BEST AVAILABLE COPY

(54) Title: CONSENSUS/ANCESTRAL IMMUNOGENS

(57) Abstract: The present invention relates, in general, to an immunogen and, in particular, to an immunogen for inducing antibod-
ies that neutralizes a wide spectrum of HIV primary isolates and/or to an immunogen that induces a T cell immune response. The
invention also relates to a method of inducing anti-HIV antibodies, and/or to a method of inducing a T cell immune response, using
such an immunogen. The invention further relates to nucleic acid sequences encoding the present immunogens.

CONSENSUS/ANCESTRAL IMMUNOGENS

This application claims priority from Prov.
Appln. No. 60/503,460, filed September 17, 2003, and
Prov. Appln. No. 60/604,722, filed August 27, 2004,
5   the entire contents of which are incorporated herein
by reference.

TECHNICAL FIELD

The present invention relates, in general, to
an immunogen and, in particular, to an immunogen for
10   inducing antibodies that neutralize a wide spectrum
of HIV primary isolates and/or to an immunogen that
induces a T cell immune response. The invention
also relates to a method of inducing anti-HIV
antibodies, and/or to a method of inducing a T cell
15   immune response, using such an immunogen. The
invention further relates to nucleic acid sequences
encoding the present immunogens.

BACKGROUND

The high level of genetic variability of HIV-1
20   has presented a major hurdle for AIDS vaccine
development. Genetic differences among HIV-1 groups
M, N, and O are extensive, ranging from 30% to 50%
in *gag* and *env* genes, respectively (Gurtler et al,
J. Virol. 68:1581-1585 (1994), Vanden Haesevelde et
25   al, J. Virol. 68:1586-1596 (1994), Simon et al, Nat.
Med. 4:1032-1037 (1998), Kuiken et al, Human

retroviruses and AIDS 2000: a compilation and
analysis of nucleic acid and amino acid sequences
(Theoretical Biology and Biophysics Group, Los
Alamos National Laboratory, Los Alamos, New
5   Mexico)). Viruses within group M are further
classified into nine genetically distinct subtypes
(A-D, F-H, J and K) (Kuiken et al, Human
retroviruses and AIDS 2000: a compilation and
analysis of nucleic acid and amino acid sequences
10   (Theoretical Biology and Biophysics Group, Los
Alamos National Laboratory, Los Alamos, New Mexico,
Robertson et al, Science 288:55-56 (2000), Robertson
et al, Human retroviruses and AIDS 1999: a
compilation and analysis of nucleic acid and amino
15   acid sequences, eds. Kuiken et al (Theoretical
Biology and Biophysics Group, Los Alamos National
Laboratory, Los Alamos, New Mexico), pp. 492-505
(2000)). With the genetic variation as high as 30%
in *env* genes among HIV-1 subtypes, it has been
20   difficult to consistently elicit cross-subtype T and
B cell immune responses against all HIV-1 subtypes.
HIV-1 also frequently recombines among different
subtypes to create circulating recombinant forms
(CRFs) (Robertson et al, Science 288:55-56 (2000),
25   Robertson et al, Human retroviruses and AIDS 1999: a
compilation and analysis of nucleic acid and amino
acid sequences, eds. Kuiken et al (Theoretical
Biology and Biophysics Group, Los Alamos National
Laboratory, Los Alamos, New Mexico), pp. 492-505
30   (2000), Carr et al, Human retroviruses and AIDS
1998: a compilation and analysis of nucleic acid and

amino acid sequences, eds. Korber et al (Theoretical

Biology and Biophysics Group, Los Alamos National

Laboratory, Los Alamos, New Mexico), pp. III-10-III-

19 (1998)). Over 20% of HIV-1 isolates are

5   recombinant in geographic areas where multiple

subtypes are common (Robertson et al, Nature

374:124-126 (1995), Cornelissen et al, J. virol.

70:8209-8212 (1996), Dowling et al, AIDS 16:1809-

1820 (2002)), and high prevalence rates of

10   recombinant viruses may further complicate the

design of experimental HIV-1 immunogens.

To overcome these challenges in AIDS vaccine

development, three computer models (consensus,

ancestor and center of the tree) have been used to

15   generate centralized HIV-1 genes to (Gaschen et al,

Science 296:2354-2360 (2002), Gao et al, Science

299:1517-1518 (2003), Nickle et al, Science

299:1515-1517 (2003), Novitsky et al, J. Virol.

76:5435-5451 (2002), Ellenberger et al, Virology

20   302:155-163 (2002), Korber et al, Science 288:1789-

1796 (2000)). The biology of HIV gives rise to

star-like phylogenies, and as a consequence of this,

the three kinds of sequences differ from each other

by 2 - 5% (Gao et al, Science 299:1517-1518 (2003)).

25   Any of the three centralized gene strategies will

reduce the protein distances between immunogens and

field virus strains. Consensus sequences minimize

the degree of sequence dissimilarity between a

vaccine strain and contemporary circulating viruses

30   by creating artificial sequences based on the most

common amino acid in each position in an alignment

(Gaschen et al, Science 296:2354-2360 (2002)).
Ancestral sequences are similar to consensus
sequences but are generated using maximum-likelihood
phylogenetic analysis methods (Gaschen et al,
5   Science 296:2354-2360 (2002), Nickle et al, Science
299:1515-1517 (2003)) . In doing so, this method
recreates the hypothetical ancestral genes of the
analyzed current wild-type sequences (Figure 26).
Nickle et al proposed another method to generate
10   centralized HIV-1 sequences, center of the tree
(COT), that is similar to ancestral sequences but
less influenced by outliers (Science 299:1515-1517
(2003)).

        The present invention results, at least in
15   part, from the results of studies designed to
determine if centralized immunogens can induce both
T and B cell immune responses in animals. These
studies involved the generation of an artificial
group M consensus env gene (CON6), and construction
20   of DNA plasmids and recombinant vaccinia viruses to
express CON6 envelopes as soluble gp120 and gp140CF
proteins. The results demonstrate that CON6 Env
proteins are biologically functional, possess
linear, conformational and glycan-dependent epitopes
25   of wild-type HIV-1,· and induce cytokine-producing T
cells that recognize T cell epitopes of both HIV
subtypes B and C. Importantly, CON6 gp120 and
gp140CF proteins induce antibodies that neutralize
subsets of subtype B and C HIV-1 primary isolates.
30   The iterative nature of study of the
centralized HIV-1 gene approach is derived from the

rapidly expanding evolution of HIV-1 sequences, and
the fact that sequences collected in the HIV
sequence database (that is, the Los Alamos National
Database) are continually being updated with new

5   sequences each year. The CON6 gp120 envelope gene
derives from Year 1999 Los Alamos National Database
sequences, and Con-S derives from Year 2000 Los
Alamos National Database sequences. In addition,
CON6 has Chinese subtype C V1, V2, V4, and V5 Env

10   sequences, while Con-S has all group M consensus Env
constant and variable regions, that have been
shortened to minimal-length variable loops. Codon-
optimized genes for a series of Year 2003 group M
and subtype consensus sequences have been designed,

15   as have a corresponding series of wild-type HIV-1
Env genes for comparison, for use in inducing
broadly reactive T and B cell responses to HIV-1
primary isolates.

SUMMARY OF THE INVENTION

20   The present invention relates to an immunogen
for inducing antibodies that neutralize a wide
spectrum of HIV primary isolates and/or to an
immunogen that induces a T cell immune response, and
to nucleic acid sequences encoding same. The

25   invention also relates to a method of inducing anti-
HIV antibodies, and/or to a method of inducing a T
cell immune response, using such an immunogen.
        Objects and advantages of the present invention
will be clear from the description that follows.

BRIEF DESCRIPTION OF THE DRAWINGS

     Figures 1A-1D: Generation and expression of
the group M consensus *env* gene (CON6). The complete
amino acid sequence of CON6 gp160 is shown.
5   (Fig. 1A) The five regions from the wild-type
CRF08_BC (98CN006) *env* gene are indicated by
underlined letters. Variable regions are indicated
by brackets above the sequences. Potential N-liked
glycosylation sites are highlighted with bold-faced
10 letters. (Fig. 1B) Constructs of CON6 gp120 and
gp140CF. CON6 gp120 and gp140CF plasmids were
engineered by introducing a stop codon after the
gp120 cleavage site or before the transmembrane
domain, respectively. The gp120/gp41 cleavage site
15 and fusion domain of gp41 were deleted in the
gp140CF protein. (Fig.1C) Expression of CON6 gp120
and gp140CF. CON6 gp120 and gp140CF were purified
from the cell culture supernatants of rVV-infected
293T cells with *galanthus Nivalis argarose* lectin
20 columns. Both gp120 and gp140CF were separated on a
10% SDS-polyarylamide gel and stained with Commassie
blue. (Fig. 1D.) CON6 *env* gene optimized based on
codon usage for highly expressed human genes.

     Figures 2A-2E. Binding of CON6 gp120 gp140 CF
25 to soluble CD4 (sCD4) and anti-Env mAbs. (Figs. 2A-
2B) Each of the indicated mabs and sCD4 was
covalently immobilized to a CM5 sensor chip
(BIAcore) and CON6 gp120 (Fig. 2A) or gp140CF (Fig.

2B) (100 µg/ml and 300 µg/ml, respectively) were
injected over each surface. Both gp120 and gp140CF
proteins reacted with each anti-gp120 mabs tested
except for17b mab, which showed negligible binding
5   to both CON6 gp120 and gp140CF. To determine
induction of 17b mab binding to CON6 gp120 and
gp140CF, CON6 gp120 (Fig. 2C) or gp140CF (Fig. 2D)
proteins were captured (400-580 RU) on individual
flow cells immobilized with sCD4 or mabs A32 or T8.
10   Following stabilization of each of the surface, mAb
17b was injected and flowed over each of the
immobilized flow cells. Overlay of curves show that
the binding of mab 17b to CON6 Env proteins was
markedly enhanced on both sCD4 and mab A32 surfaces
15   but not on the T8 surface (Figs. 2C-2D). To
determine binding of CON6 gp120 and gp140CF to human
mabs in ELISA, stock solutions of 20µg/ml of mabs
447, F39F, A32, IgG1b12 and 2F5 on CON6 gp120 and
gp140CF were tittered (Fig. 2E). Mabs 447 (V3),
20   F39F (V3) A32 (gp120) and IgG1b12 (CD4 binding site)
each bound to both CON6 gp120 and 140 well, while
2F5 (anti-gp41 ELDKWAS) only bound gp140CF. The
concentration at endpoint titer on gp120 for mab 447
and F39F binding was <0.003 µg/ml and 0.006 µg/ml,
25   respectively; for mab A32 was <0.125 µg/ml; for
IgG1b12 was <0.002 µg/ml; and for 2F5 was 0.016
µg/ml.

Figures 3A and 3B. Infectivity and coreceptor
usage of CON6 envelope. (Fig. 3A) CON6 and control

env plasmids were cotransfected with HIV-1/SG3Δenv
backbone into human 293T cells to generate Env-
pseudovirions. Equal amounts of each pseudovirion
(5 ng p24) were used to infect JC53-BL cells. The

5   infectivity was determined by counting the number of
blue cells (infectious units, IU) per microgram of
p24 of pseudovirons (IU/μg p24) after staining the
infected cells for β-gal expression. (Fig. 3B)
Coreceptor usage of the CON6 *env* gene was determined

10   on JC53BL cells treated with AMD3100 and/or TAK-799
for 1 hr (37°C) then infected with equal amounts of
p24 (5 ng) of each Env-pseudovirion. Infectivity in
the control group (no blocking agent) was set as
100%. Blocking efficiency was expressed as the

15   percentage of IU from blocking experiments compared
to those from control cultures without blocking
agents. Data shown are mean ± SD.

     Figure 4. Western blot analysis of multiple
subtype Env proteins against multiple subtype

20   antisera. Equal amount of Env proteins (100 ng)
were separated on 10% SDS-polyacrylamide gels.
Following electrophoresis, proteins were transferred
to Hybond ECL nitrocellulose membranes and reacted
with sera from HIV-1 infected patients (1:1,000) or

25   guinea pigs immunized with CON6 gp120 DNA prime, rVV
boost (1:1,000). Protein-bound antibody was probed
with fluorescent-labeled secondary antibodies and
the images scanned and recorded on an infrared
imager Odyssey (Li-Cor, Lincoln, NE). Subtypes are

indicated by single-letters after Env protein and
serum IDs. Four to six sera were tested for each
subtype, and reaction patterns were similar among
all sera from the same subtype. One representative
5   result for each subtype serum is shown.

    Figure 5.   T cell immune responses induced by
CON6 Env immunogens in mice. Splenocytes were
isolated from individual immunized mice (5
mice/group). After splenocytes were stimulated *in*
10   *vitro* with overlapping Env peptide pools of CON6
(black column), subtype B (hatched column), subtype
C (white column), and medium (no peptide; gray
column), INF-γ producing cells were determined by
the ELISPOT assay. T cell IFN-γ responses induced
15   by either CON6 gp120 or gp140CF were compared to
those induced by subtype specific Env immunogens
(JRFL and 96ZM651). Total responses for each
envelope peptide pool are expressed as SFCs per
million splenocytes. The values for each column are
20   the mean ± SEM of IFN-γ SFCs (n=5 mice/group).

    Figures 6A-6E. Construction of codon usage
optimized subtype C ancestral and consensus envelope
genes (Figs. 6A and 6B, respectively). Ancestral
and consensus amino acid sequences (Figs. 6C and 6D,
25   respectively) were transcribed to mirror the codon
usage of highly expressed human genes. Paired
oligonucleotides (80-mers) overlapping by 20 bp were
designed to contain 5' invariant sequences including

the restriction enzyme sites EcoRI, BbsI, Bam HI and BsmBI. BbsI and BsmBI are Type II restriction enzymes that cleave outside of their recognition sequences. Paired oligomers were linked

5   individually using PCR and primers complimentary to the 18 bp invariant sequences in a stepwise fashion, yielding 140bp PCR products. These were subcloned into pGEM-T and sequenced to confirm the absence of inadvertant mutations/deletions. Four individual

10   pGEM-T subclones containing the proper inserts were digested and ligated together into pcDNA3.1. Multi-fragment ligations occurred repeatly amongst groups of fragments in a stepwise manner from the 5' to the 3' end of the gene until the entire gene was

15   reconstructed in pcDNA3.1. (See schematic in Fig. 6E.)

     Figure 7. JC53-BL cells are a derivative of HeLa cells that express high levels of CD4 and the HIV-1 coreceptors CCR5 and CXCR4. They also contain

20   the reporter cassettes of luciferase and β-galactosidase that are each expressed from an HIV-1 LTR. Expression of the reporter genes is dependent on production of HIV-1 Tat. Briefly, cells are seeded into 24 or 96-well plates, incubated at 37°C

25   for 24 hours and treated with DEAE-Dextran at 37°C for 30 minutes. Virus is serially diluted in 1% DMEM, added to the cells incubating in DEAE-Dextran, and allowed to incubate for 3 hours at 37°C after which an additional cell media is added to each

well. Following a final 48-hour incubation at 37°C,
cells are either fixed, stained using X-Gal to
visualize β-galactosidase expressing blue foci or
frozen-thawed three times to measure luciferase
5   activity.

        Figure 8.   Sequence alignment of subtype C
ancestral and consensus *env* genes. Alignment of the
subtype C ancestral (bottom line) and consensus (top
line) *env* sequences showing a 95.5% sequence
10   homology; amino acid sequence differences are
indicated. One noted difference is the addition of a
glycosylation site in the C ancestral *env* gene at
the base of the V1 loop. A plus sign indicates a
within-class difference of amino acid at the
15   indicated position; a bar indicates a change in the
class of amino acid. Potential N-glycosylation sites
are marked in blue. The position of truncation for
the gp140 gene is also shown.

        Figure 9.   Expression of subtype C ancestral
20   and consensus envelopes in 293T cells. Plasmids
containing codon-optimized *gp160, gp140, or gp120*
subtype C ancestral and consensus genes were
transfected into 293T cells, and protein expression
was examined by Western Blot analysis of cell
25   lysates. 48-hours post-transfection, cell lysates
were collected, total protein content determined by
the BCA protein assay, and 2 μg of total protein was
loaded per lane on a 4-20% SDS-PAGE gel. Proteins

were transferred to a PVDF membrane and probed with
HIV-1 plasma from a subtype C infected patient.

    Figures 10A and 10B. Fig. 10A. *Trans*
complementation of *env*-deficient HIV-1 with codon-
5   optimized subtype C ancestral and consensus gp160
and gp140. Plasmids containing codon-optimized,
subtype C ancestral or consensus *gp160* or *gp140*
genes were co-transfected into 293T cells with an
HIV-1/SG3Δ*env* provirus. 48 hours post-transfection
10 cell supernatants containing pseudotyped virus were
harvested, clarified by centrifugation, filtered
through at 0.2μM filter, and pelleted through a 20%
sucrose cushion. Quantification of p24 in each
virus pellet was determined using the Coulter HIV-1
15 p24 antigen assay; 25ng of p24 was loaded per lane
on a 4-20% SDS-PAGE gel for particles containing a
codon-optimized envelope. 250ng of p24 was loaded
per lane for particles generated by co-transfection
of a rev-dependent wild-type subtype C 96ZAM651*env*
20 gene. Differences in the amount of p24 loaded per
lane were necessary to ensure visualization of the
rev-dependent envelopes by Western Blot. Proteins
were transferred to a PVDF membrane and probed with
pooled plasma from HIV-1 subtype B and subtype C
25 infected individuals. Fig. 10B. Infectivity of
virus particles containing subtype C ancestral and
consensus envelope glycoproteins. Infectivity of
pseudotyped virus containing ancestral or consensus
gp160 or *gp140* envelope was determined using the

JC53-BL assay. Sucrose cushion purified virus
particles were assayed by the Coulter p24 antigen
assay, and 5-fold serial dilutions of each pellet
were incubated with DEAE-Dextran treated JC53-BL
cells. Following a 48-hour incubation period, cells
were fixed and stained to visualize β-galactosidase
expressing cells. Infectivity is represented as
infectious units per ng of p24 to normalize for
differences in the concentration of the input
pseudovirions.

Figure 11. Co-receptor usage of subtype C
ancestral and consensus envelopes. Pseudotyped
particles containing ancestral or consensus envelope
were incubated with DEAE-Dextran treated JC53-BL
cells in the presence of AMD3100 (a specific
inhibitor of CXCR4), TAK779 (a specific inhibitor of
CCR5), or AMD3000+TAK779 to determine co-receptor
usage. NL4.3, an isolate known to utilize CXCR4,
and YU-2, a known CCR5-using isolate, were included
as controls.

Figures 12A-12C. Neutralization sensitivity of
subtype C ancestral and consensus envelope
glycoproteins. Equivalent amounts of pseudovirions
containing the ancestral, consensus or 96ZAM651
gp160 envelopes (1,500 infectious units) were pre-
incubated with a panel of plasma samples from HIV-1
subtype C infected patients and then added to the
JC53-BL cell monolayer in 96-well plates. Plates

were cultured for two days and luciferase activity
was measured as an indicator of viral infectivity.
Virus infectivity is calculated by dividing the
luciferase units (LU) produced at each concentration

5   of antibody by the LU produced by the control
infection. The mean 50% inhibitory concentration
(IC_{50}) and the actual % neutralization at each
antibody dilution are then calculated for each
virus. The results of all luciferase experiments

10   are confirmed by direct counting of blue foci in
parallel infections.

        Figures 13A-13F. Protein expression of
consensus subtype C Gag (Fig. 13A) and Nef (Fig.
13B) following transfection into 293T cells.

15   Consensus subtype C Gag and Nef amino acid sequences
are set forth in Figs. 13C and 13D, respectively,
and encoding sequences are set forth in Figs. 13E
and 13F, respectively.

        Figures 14A-14C. Figs. 14A and 14B show the

20   Con-S Env amino acid sequence and encoding sequence,
respectively. Fig. 14C shows expression of Group M
consensus Con-S Env proteins using an *in vitro*
transcription and translation system.

        Figures 15A and 15B. Expression of Con-S *env*

25   gene in mammalian cells. (Fig. 15A – cell lysate,
Fig. 15B – supernatant.)

                                 14

Figures 16A and 16B. Infectivity (Fig. 16A)
and coreceptor usage (Fig. 16B) of CON6 and Con-S
env genes.

Figures 17A-17C. Env protein incorporation in
5   CON6 and Con-S Env-pseudovirions. (Fig. 17A –
lysate, Fig. 17B – supernatant, Fig. 17C pellet.)

Figures 18A-18D. Figs. 18A and 18B show
subtype A consensus Env amino acid sequence and
nucleic acid sequence encoding same, respectively.
10   Figs. 18C and 18D show expression of A.con *env* gene
in mammalian cells (Fig. 18C – cell lysate, Fig. 18D
– supernatant).

Figures 19A-19H. M.con.gag (Fig. 19A),
M.con.pol (Fig. 19B), M.con.nef (Fig. 19C) and
15   C.con.pol (Fig. 19D) nucleic acid sequences and
corresponding encoded amino acid sequences (Figs.
19E-19H, respectively).

Figures 20A-20D. Subtype B consensus *gag* (Fig.
20A) and *env* (Fig.20B) genes. Corresponding amino
20   acid sequences are shown in Figs. 20C and 20D.

Figure 21. Expression of subtype B consensus
env and *gag* genes in 293T cells. Plasmids
containing codon-optimized subtype B consensus
gp160, *gp140*, and *gag* genes were transfected into
25   293T cells, and protein expression was examined by

Western Blot analysis of cell lysates. 48-hours
post-transfection, cell lysates were collected,
total protein content determined by the BCA protein
assay, and 2 μg of total protein was loaded per lane
5   on a 4-20% SDS-PAGE gel.   Proteins were transferred
to a PVDF membrane and probed with serum from an
HIV-1 subtype B infected individual.

    Figure 22.   Co-receptor usage of subtype B
consensus envelopes.   Pseudotyped particles
10 containing the subtype B consensus gp160 Env were
incubated with DEAE-Dextran treated JC53-BL cells in
the presence of AMD3100 (a specific inhibitor of
CXCR4), TAK779 (a specific inhibitor of CCR5), and
AMD3000+TAK779 to determine co-receptor usage.
15 NL4.3, an isolate known to utilize CXCR4 and YU-2, a
known CCR5-using isolate, were included as controls.

    Figures 23A and 23B.   *Trans* complementation of
env-deficient HIV-1 with codon-optimized subtype B
consensus *gp160* and *gp140* genes.   Plasmids
20 containing codon-optimized, subtype B consensus
gp160 or *gp140* genes were co-transfected into 293T
cells with an HIV-1/SG3Δenv provirus.   48-hours
post-transfection cell supernatants containing
pseudotyped virus were harvested, clarified in a
25 tabletop centrifuge, filtered through a 0.2μM
filter, and pellet through a 20% sucrose cushion.
Quantification of p24 in each virus pellet was
determined using the Coulter HIV-1 p24 antigen

16

assay; 25 ng of p24 was loaded per lane on a 4-20%
SDS-PAGE gel.   Proteins were transferred to a PVDF.
membrane and probed with anti-HIV-1 antibodies from
infected HIV-1 subtype B patient serum.   *Trans*
5   complementation with a rev-dependent NL4.3 *env* was
included for control.   Figure 23B.   Infectivity of
virus particles containing the subtype B concensus
envelope.   Infectivitiy of pseudotyped virus
containing consensus B gp160 or gp140 was determined
10   using the JC53-BL assay.   Sucrose cushion purified
virus particles were assayed by the Coulter p24
antigen assay, and 5-fold serial dilutions of each
pellet were incubated with DEAE-Dextran treated
JC53-BL cells.   Following a 48-hour incubation
15   period, cells were fixed and stained to visualize β-
galactosidase expressing cells.   Infectivity is
expressed as infectious units per ng of p24.

Figures 24A-24D.   Neutralization sensitivity of
virions containing subtype B consensus gp160
20   envelope.   Equivalent amounts of pseudovirions
containing the subtype B consensus or NL4.3 Env
(gp160) (1,500 infectious units) were preincubated
with three different monoclonal neutralizing
antibodies and a panel of plasma samples from HIV-1
25   wubtype B infected individuals, and then added to
the JC53-BL cell monolayer in 96-well plates.
Plates were cultured for two days and luciferase
activity was measured as an indicator of viral
infectivity.   Virus infectivity was calculated by

dividing the luciferase units (LU) produced at each
concentration of antibody by the LU produced by the
control infection. The mean 50% inhibitory
concentration (IC_{50}) and the actual % neutralization
5   at each antibody dilution were then calculated for
each virus. The results of all luciferase
experiments were confirmed by direct counting of
blue foci in parallel infections. Fig. 24A.
Neutralization of Pseudovirions containing Subtype B
10   consensus Env (gp160). Fig. 24B. Neutralization of
Pseudovirions containing NL4.3 Env (gp160).
Fig. 24C. Neutralization of Pseudovirions containing
Subtype B consensus Env (gp160). Fig. 24D.
Neutralization of Pseudovirions containing NL4.3 Env
15   (gp160).

Figures 25A and 25B. Fig. 25A. Density and p24
analysis of sucrose gradient fractions. 0.5ml
fractions were collected from a 20-60% sucrose
gradient. Fraction number 1 represents the most
20   dense fraction taken from the bottom of the gradient
tube. Density was measured with a refractometer and
the amount of p24 in each fraction was determined by
the Coulter p24 antigen assay. Fractions 6-9, 10-
15, 16-21, and 22-25 were pooled together and
25   analyzed by Western Blot. As expected, virions
sedimented at a density of 1.16-1.18 g/ml.
Fig. 25B. VLP production by co-transfection of
subtype B consensus *gag* and *env* genes. 293T cells
were co-transfected with subtype B consensus *gag* and

env genes. Cell supernatants were harvested 48-
hours post-transfection, clarified through at 20%
sucrose cushion, and further purified through a 20-
60% sucrose gradient. Select fractions from the
5   gradient were pooled, added to 20ml of PBS, and
centrifuged overnight at 100,000 x g. Resuspended
pellets were loaded onto a 4-20% SDS-PAGE gel,
proteins were transferred to a PVDF membrane, and
probed with plasma from an HIV-1 subtype B infected
10 individual.

      Figures 26A and 26B. Fig. 26A. 2000 Con-S
140CFI.ENV. Fig. 26B. Codon-optimized Year 2000
Con-S 140CFI.seq.

      Figure 27. Individual C57BL/6 mouse T cell
15 responses to HIV-1 envelope peptides. Comparative
immunogenicity of CON6 gp140CFI and Con-S gp140CFI
in C57BL/C mice. Mice were immunized with either
HIV5305 (Subtype A), 2801 (Subtype B), CON6 or Con-S
Envelope genes in DNA prime, rVV boost regimens, 5
20 mice per group. Spleen cells were assayed for IFN-γ
spot-forming cells 10 days after rVV boost, using
mixtures of overlapping peptides from Envs of HIV-1
UG37(A), MN(B), Ch19(C), 89.6(B) SF162(B) or no
peptide negative control.

25   Figures 28A-28C. Fig. 28A. Con-B 2003 Env. pep
(841 a.a.). Amino acid sequence underlined is the
fusion domain that is deleted in 140CF design and

the "W" underlined is the last amino acid at the
C-terminus, all amino acids after the "W" are
deleted in the 140CF design. Fig. 28B. Con-B-
140CF.pep (632 a.a.). Amino acids in bold identify

5   the junction of the deleted fusion cleavage site.
Fig. 28C. Codon-optimized Con-B 140CF.seq
(1927 nt.).

       Figures 29A-29C. Fig. 29A. CON_OF_CONS-2003
(829 a.a.). Amino acid sequence underlined is the

10   fusion domain that is deleted in 140CF design and
the "W" underlined is the last amino acid at the
C-terminus, all amino acids after the "W" are
deleted in the 140CF design. Fig. 29B. ConS-2003
140CF.pep (620 a.a.). Amino acids in bold identify

15   the junction of the deleted fusion cleavage site.
Fig. 29C. CODON-OPTIMIZED ConS-2003 140CF.seq (1891
nt.).

       Figures 30A-30C. Fig. 30A. CONSENSUS_A1-2003
(845 a.a.). Amino acid sequence underlined is the

20   fusion domain that is deleted in 140CF design and
the "W" underlined is the last amino acid at the
C-terminus, all amino acids after the "W" are
deleted in the 140CF design. Fig. 30B. Con-A1-2003
140CF.pep (629 a.a.). Amino acids in bold identify

25   the junction of the deleted fusion cleavage site.
Fig. 30C. CODON-OPTIMIZED Con-A1-2003.seq.

Figures 31A-31C. Fig. 31A. CONSENSUS_C-2003

(835 a.a.). Amino acid sequence underlined is the

fusion domain that is deleted in 140CF design and

the "W" underlined is the last amino acid at the

5   C-terminus, all amino acids after the "W" are

deleted in the 140CF design. Fig. 31B. Con-C 2003

140CF.pep (619 a.a.). Amino acids in bold identify

the junction of the deleted fusion cleavage site.

Fig. 31C. CODON-OPTIMIZED Con-C-2003 (140 CF (1,888

10   nt.).

Figures 32A-32C. Fig. 32A. CONSENSUS_G-2003

(842 a.a.). Amino acid sequence underlined is the

fusion domain that is deleted in 140CF design and

the "W" underlined is the last amino acid at the

15   C-terminus, all amino acids after the "W" are

deleted in the 140CF design. Fig. 32B. Con-G-2003

140CF.pep (626 a.a.). Amino acids in bold identify

the junction of the deleted fusion cleavage site.

Fig. 32C. CODON-OPTIMIZED Con-G-2003.seq.

20   Figures 33A-33C. Fig. 33A. CONSENSUS_01_AE-

2003 (854 a.a.). Amino acid sequence underlined is

the fusion domain that is deleted in 140CF design

and the "W" underlined is the last amino acid at

the C-terminus, all amino acids after the "W" are

25   deleted in the 140CF design. Fig. 33B. Con-AE01-

2003 140CF.pep (638 a.a.). Amino acids in bold

identify the junction of the deleted fusion cleavage

site. Fig. 33C. CODON-OPTIMIZED Con-AE01-2003.seq.
(1945 nt.).

Figures 34A-34C. Fig. 34A. Wild-type subtype
A Env. 00KE_MSA4076-A (Subtype A, 891 a.a.). Amino
5   acid sequence underlined is the fusion domain that
is deleted in 140CF design and the "W" underlined
is the last amino acid at the C-terminus, all amino
acids after the "W" are deleted in the 140CF design.
Fig. 34B. 00KE_MSA4076-A 140CF.pep (647 a.a.).
10 Amino acids in bold identify the junction of the
deleted fusion cleavage site. Fig. 34C. CODON-
OPTIMIZED 00KE_MSA4076-A 140CF.seq. (1972 nt.).

Figures 35A-35C. Fig. 35A. Wild-type subtype
B. QH0515.1g gp160 (861 a.a.). Amino acid sequence
15 underlined is the fusion domain that is deleted in
140CF design and the "W" underlined is the last
amino acid at the C-terminus, all amino acids after
the "W" are deleted in the 140CF design. Fig. 35B.
QH0515.1g 140CF (651 a.a.). Amino acids in bold
20 identify the junction of the deleted fusion cleavage
site. Fig. 35C. CODON-OPTIMIZED QH0515.1g
140CF.seq (1984 nt.).

Figures 36A-36C. Fig. 36A. Wild-type subtype
C. DU123.6 gp160 (854 a.a.). Amino acid sequence
25 underlined is the fusion domain that is deleted in
140CF design and the "W" underlined is the last
amino acid at the C-terminus, all amino acids after

the "W" are deleted in the 140CF design. Fig. 36B.
DU123.6 140CF (638 a.a.). Amino acids in bold
identify the junction of the deleted fusion cleavage
site. Fig. 36C. CODON-OPTIMIZED DU123.6 140CF.seq
5   (1945 nt.).

      Figures 37A-37C. Fig. 37A. Wild-type subtype
CRF01_AE. 97CNGX2F-AE (854 a.a.). Amino acid
sequence underlined is the fusion domain that is
deleted in 140CF design and the "W" underlined is
10   the last amino acid at the C-terminus, all amino
acids after the "W" are deleted in the 140CF design.
Fig. 37B. 97CNGX2F-AE 140CF.pep (629 a.a.). Amino
acids in bold identify the junction of the deleted
fusion cleavage site. Fig. 37C. CODON-OPTIMIZED
15   97CNGX2F-AE 140CF.seq (1921 nt.).

      Figures 38A-38C. Fig. 38A. Wild-type DRCBL-G
(854 a.a.). Amino acid sequence underlined is the
fusion domain that is deleted in 140CF design and
the "W" underlined is the last amino acid at the
20   C-terminus, all amino acids after the "W" are
deleted in the 140CF design. Fig. 38B. DRCBL-G
140CF.pep (630 a.a.). Amino acids in bold identify
the junction of the deleted fusion cleavage site.
Fig. 38C. CODON-OPTIMIZED DRCBL-G 140CF.seq (1921
25   nt.).

Figures 39A and 39B.   Fig. 39A.   2003 Con-S
Env. Fig. 39B.   2003 Con-S Env.seq.opt.
(Seq.opt. = codon optimized encoding sequence.)

Figures 40A and 40B.   Fig. 40A.   2003 M.
5   Group.Anc Env.   Fig. 40B.   2003 M. Group.anc
Env.seq.opt.   (Seq.opt. = codon optimized encoding
sequence.)

Figures 41A and 41B.   Fig. 41A.   2003 CON_A1
Env.   Fig. 41B.   2003 CON_A1 Env.seq.opt.
10   (Seq.opt. = codon optimized encoding sequence.)

Figures 42A and 42B.   Fig. 42A.   2003 A1.Anc
Env.   Figs. 42B.   2003 A1.anc Env.seq.opt.
(Seq.opt. = codon optimized encoding sequence.)

Figures 43A and 43B.   Fig. 43A.   2003 CON_A2
15   Env.   Fig. 43B.   2003 CON_A2 Env.seq.opt.
(Seq.opt. = codon optimized encoding sequence.)

Figures 44A and 44B.   Fig. 44A.   2003 CON_B
Env.   Fig. 44B.   2003 CON_B Env.seq.opt.
(Seq.opt. = codon optimized encoding sequence.)

20   Figures 45A and 45B.   Fig. 45A.   2003 B.anc
Env.   Figs. 45B.   2003 B.anc Env.seq.opt.
(Seq.opt. = codon optimized encoding sequence.)

Figures 46A and 46B.   Fig. 46A.   2003 CON_C
Env.   Fig. 46B.   2003 CON_C Env.seq.opt.
(Seq.opt. = codon optimized encoding sequence.)

Figures 47A and 47B.   Fig. 47A.   2003 C.anc
5   Env.   Fig. 47B.   2003 C.anc Env.seq.opt.
(Seq.opt. = codon optimized encoding sequence.)

Figures 48A and 48B.   Fig. 48A.   2003 CON_D
Env.   Fig. 48B.   2003 CON_D Env.seq.opt.
(Seq.opt. = codon optimized encoding sequence.)

10   Figures 49A and 49B.   Fig. 49A.   2003 CON_F1
Env.   Fig. 49B.   2003 CON_F1 Env.seq.opt.
(Seq.opt. = codon optimized encoding sequence.)

Figures 50A and 50B.   Fig. 50A.   2003 CON_F2
Env.   Fig. 50B.   2003 CON_F2 Env.seq.opt.
15   (Seq.opt. = codon optimized encoding sequence.)

Figures 51A and 51B.   Fig. 51A.   2003 CON_G
Env.   Fig. 51B.   2003 CON_G Env.seq.opt.
(Seq.opt. = codon optimized encoding sequence.)

Figures 52A and 52B.   Fig. 52A.   2003 CON_H
20   Env.   Fig. 52B.   2003 CON_H Env.seq.opt.
(Seq.opt. = codon optimized encoding sequence.)

Figures 53A and 53B. Fig. 53A. 2003 CON_01_AE Env. Fig. 53B. 2003 CON_01_AE Env.seq.opt. (Seq.opt. = codon optimized encoding sequence.)

Figures 54A and 54B. Fig. 54A. 2003 CON_02_AG Env. Fig. 54B. 2003 CON_02_AG Env.seq.opt. (Seq.opt. = codon optimized encoding sequence.)

Figures 55A and 55B. Fig. 55A. 2003 CON_03_AB Env. Fig. 55B. 2003 CON_03_AB Env.seq.opt. (Seq.opt. = codon optimized encoding sequence.)

Figures 56A and 56B. Fig. 56A. 2003 CON_04_CPX Env. Fig. 56B. 2003 CON_04_CPX Env.seq.opt. (Seq.opt. = codon optimized encoding sequence.)

Figures 57A and 57B. Fig. 57A. 2003 CON_06_CPX Env. Fig. 57B. 2003 CON_06_CPX Env.seq.opt. (Seq.opt. = codon optimized encoding sequence.)

Figures 58A and 58B. Fig. 58A. 2003 CON_08_BC Env. Fig. 58B. 2003 CON_08_BC Env.seq.opt. (Seq.opt. = codon optimized encoding sequence.)

Figures 59A and 59B. Fig. 59A. 2003 CON_10_CD Env. Fig. 59B. 2003 CON_10_CD Env.seq.opt. (Seq.opt. = codon optimized encoding sequence.)

Figures 60A and 60B. Fig. 60A. 2003
CON_11_CPX Env. Fig. 60B. 2003 CON_11_CPX
Env.seq.opt. (Seq.opt. = codon optimized encoding
sequence.)

5   Figures 61A and 61B. Fig. 61A. 2003 CON_12_BF
Env. Fig. 61B. 2003 CON_12_BF Env.seq.opt.
(Seq.opt. = codon optimized encoding sequence.)

Figures 62A and 62B. Fig. 62A. 2003 CON_14_BG
Env. Fig. 62B. 2003 CON_14_BG Env.seq.opt.
10   (Seq.opt. = codon optimized encoding sequence.)

Figures 63A and 63B. Fig. 63A. 2003_CON_S
gag.PEP. Fig. 63B. 2003_CON_S gag.OPT.
(OPT = codon optimized encoding sequence.)

Figures 64A and 64B. Fig. 64A.
15   2003_M.GROUP.anc gag.PEP. Fig. 64B.
2003_M.GROUP.anc gag.OPT. (OPT = codon optimized
encoding sequence.)

Figures 65A-65D. Fig. 65A. 2003_CON_A1
gag.PEP. Fig. 65B. 2003_CON_A1 gag.OPT. Fig. 65C.
20   2003_A1.anc gag.PEP. Fig. 65D. 2003_A1.anc
gag.OPT. (OPT = codon optimized encoding sequence.)

Figures 66A and 66B. Fig. 66A. 2003_CON_A2
gag.PEP. Fig. 66B. 2003_CON_A2 gag.OPT.
(OPT = codon optimized encoding sequence.)

Figures 67A-67D. Fig. 67A. 2003_CON_B
5   gag.PEP. Fig. 67B. 2003_CON_B gag.OPT. Fig. 67C.
2003_B.anc gag.PEP. Fig. 67D. 2003_B.anc gag.OPT.
(OPT = codon optimized encoding sequence.)

Figures 68A-68D. Fig. 68A. 2003_CON_C
gag.PEP. Fig. 68B. 2003_CON_C gag.OPT. Fig. 68C.
10   2003_C.anc.gag.PEP. Fig. 68D. 2003_C.anc.gag.OPT.
(OPT = codon optimized encoding sequence.)

Figures 69A and 69B. Fig. 69A. 2003_CON_D
gag.PEP. Fig. 69B. 2003_CON_D gag.OPT.
(OPT = codon optimized encoding sequence.)

15   Figures 70A and 70B. Fig. 70A. 2003_CON_F
gag.PEP. Fig. 70B. 2003_CON_F gag.OPT.
(OPT = codon optimized encoding sequence.)

Figures 71A and 71B. Fig. 71A. 2003_CON_G
gag.PEP. Fig. 71B. 2003_CON_G gag.OPT.
20   (OPT = codon optimized encoding sequence.)

Figures 72A and 72B. Fig. 72A. 2003_CON_H
gag.PEP. Fig. 72B. 2003_CON_H gag.OPT.
(OPT = codon optimized encoding sequence.)

Figures 73A and 73B. Fig. 73A. 2003_CON_K gag.PEP. Fig. 73B. 2003_CON_K gag.OPT. (OPT = codon optimized encoding sequence.)

Figures 74A and 74B. Fig. 74A. 2003_CON_01_AE gag.PEP. Fig. 7B. 2003_CON_01_AE gag.OPT. (OPT = codon optimized encoding sequence.)

Figures 75A and 75B. Fig. 75A. 2003_CON_02_AG gag.PEP. Fig. 75B. 2003_CON_02_AG gag.OPT. (OPT = codon optimized encoding sequence.)

Figures 76A and 76B. Fig. 76A. 2003_CON_03_ABG gag.PEP. Fig. 76B. 2003_CON_03_ABG gag.OPT. (OPT = codon optimized encoding sequence.)

Figures 77A and 77B. Fig. 77A. 2003_CON_04_CFX gag.PEP. Fig. 77B. 2003 CON_04_CFX gag.OPT. (OPT = codon optimized encoding sequence.)

Figures 78A and 78B. Fig. 78A. 2003_CON_06_CPX gag.PEP. Fig. 78B. 2003_CON_06_CPX gag.OPT. (OPT = codon optimized encoding sequence.)

Figures 79A and 79B. Fig. 79A. 2003_CON_07_BC gag.PEP. Fig. 79B. 2003_CON_07_BC gag.OPT. (OPT = codon optimized encoding sequence.)

Figures 80A and 80B. Fig. 80A. 2003_CON_08_BC gag.PEP. Fig. 80B. 2003_CON_08_BC gag.OPT. (OPT = codon optimized encoding sequence.)

5   Figures 81A and 81B. Fig. 81A. 2003_CON_10_CD gag.PEP. Fig. 81B. 2003_CON_10_CD gag.OPT. (OPT = codon optimized encoding sequence.)

Figures 82A and 82B. Fig. 82A. 2003_CON_11_CPX gag.PEP. Fig. 82B. 2003_CON_11_CPX gag.OPT. (OPT = codon optimized encoding sequence.)

10   Figures 83A and 83B. Fig. 83A. 2003_CON_12_BF.gag.PEP. Fig. 83B. 2003_CON_12_BF.gag.OPT. (OPT = codon optimized encoding sequence.)

Figures 84A and 84B. Fig. 84A. 2003_CON_14_BG gag.PEP. Fig. 84B. 2003_CON_14_BG gag.OPT.
15   (OPT = codon optimized encoding sequence.)

Figures 85A and 85B. Fig. 85A. 2003_CONS nef.PEP. Fig. 85B. 2003_CONS nef.OPT. (OPT = codon optimized encoding sequence.)

20   Figures 86A and 86B. Fig. 86A. 2003_M GROUP.anc nef.PEP. Fig. 86B. 2003_M GROUP.anc.nef.OPT. (OPT = codon optimized encoding sequence.)

Figures 87A and 87B. Fig. 87A. 2003_CON_A
nef.PEP. Fig. 87B. 2003_CON_A nef.OPT.
(OPT = codon optimized encoding sequence.)

Figures 88A-88D. Fig. 88A. 2003_CON_A1
5   nef.PEP. Fig. 88B. 2003_CON_A1 nef.OPT. Fig. 88C.
2003_A1.anc nef.PEP. Fig. 88D. 2003_A1.anc
nef.OPT. (OPT = codon optimized encoding sequence.)

Figures 89A and 89B. Fig. 89A. 2003_CON_A2
nef.PEP. Fig. 89B. 2003_CON_A2 nef.OPT.
10   (OPT = codon optimized encoding sequence.)

Figures 90A-90D. Fig. 90A. 2003_CON_B
nef.PEP. Fig. 90B. 2003_CON-B nef.OPT. Fig. 90C.
2003_B.anc nef.PEP. Fig. 90D. 2003_B.anc nef.OPT.
(OPT = codon optimized encoding sequence.)

15   Figures 91A and 91B. Fig. 91A. 2003_CON_02_AG
nef.PEP. Fig. 91B. 2003_CON_02_AG nef.OPT.
(OPT = codon optimized encoding sequence.)

Figures 92A-92D. Fig. 92A. 2003_CON_C
nef.PEP. Fig. 92B. 2003_CON_C nef.OPT. Fig. 92C.
20   2003_C.anc nef.PEP. Fig. 92D. 2003_C.anc nef.OPT.
(OPT = codon optimized encoding sequence.)

Figures 93A and 93B. Fig. 93A. 2003_CON_D
nef.PEP. Fig. 93B. 2003_CON_D nef.OPT.
(OPT = codon optimized encoding sequence.)

Figures 94A and 94B. Fig. 94A. 2003_CON_F1
5   nef.PEP. Fig. 94B. 2003_CON_F1 nef.OPT.
(OPT = codon optimized encoding sequence.)

Figures 95A and 95B. Fig. 95A. 2003_CON_F2
nef.PEP. Fig. 95B. 2003_CON_F2 nef.OPT.
(OPT = codon optimized encoding sequence.)

10   Figures 96A and 96B. Fig. 96A. 2003_CON_G
nef.PEP. Fig. 96B. 2003_CON_G nef.OPT.
(OPT = codon optimized encoding sequence.)

Figures 97A and 97B. Fig. 97A. 2003_CON_H
nef.PEP. Fig. 97B. 2003_CON_H nef.OPT.
15   (OPT = codon optimized encoding sequence.)

Figures 98A and 98B. Fig. 98A. 2003_CON_01_AE
nef.PEP. Fig. 98B. 2003_CON_01_AE nef.OPT.
(OPT = codon optimized encoding sequence.)

Figures 99A and 99B. Fig. 99A. 2003_CON_03_AE
20   nef.PEP. Fig. 99B. 2003_CON_03_AE nef.OPT.
(OPT = codon optimized encoding sequence.)

Figures 100A and 100B. Fig. 100A.
2003_CON_04_CFX nef.PEP. Fig. 100B.
2003_CON_04_CFX nef.OPT. (OPT = codon optimized
encoding sequence.)

5　　　　Figures 101A and 101B. Fig. 101A.
2003_CON_06_CFX nef.PEP. Fig. 101B.
2003_CON_06_CFX nef.OPT. (OPT = codon optimized
encoding sequence.)

Figures 102A and 102B. Fig. 102A.
10　2003_CON_08_BC nef.PEP. Fig. 102B. 2003_CON_08_BC
nef.OPT. (OPT = codon optimized encoding sequence.)

Figures 103A and 103B. Fig. 103A.
2003_CON_10_CD nef.PEP. Fig. 103B. 2003_CON_10_CD
nef.OPT. (OPT = codon optimized encoding sequence.)

15　　　　Figures 104A and 104B. Fig. 104A.
2003_CON_11_CFX nef.PEP. Fig. 104B.
2003_CON_11_CFX nef.OPT. (OPT = codon optimized
encoding sequence.)

Figures 105A and 105B. Fig. 105A.
20　2003_CON_12_BF nef.PEP. Fig. 105B. 2003_CON_12_BF
nef.OPT. (OPT = codon optimized encoding sequence.)

Figures 106A and 106B. Fig. 106A.
2003_CON_14_BG nef.PEP. Fig. 106B. 2003_CON_14_BG
nef.OPT. (OPT = codon optimized encoding sequence.)

Figures 107A and 107B. Fig. 107A. 2003_CON_S
5   pol.PEP. Fig. 107B. 2003_CON_S pol.OPT.
(OPT = codon optimized encoding sequence.)

Figures 108A and 108B. Fig. 108A. 2003_M
GROUP anc pol.PEP. Fig. 108B. 2003_M.GROUP anc
pol.OPT. (OPT = codon optimized encoding sequence.)

10   Figures 109A-109D. Fig. 109A. 2003_CON_A1
pol.PEP. Fig. 109B. 2003_CON_A1 pol.OPT.
Fig. 109C. 2003_A1.anc pol.PEP. Fig. 109D.
2003_A1.anc pol.OPT. (OPT = codon optimized
encoding sequence.)

15   Figures 110A and 110B. Fig. 110A. 2003_CON_A2
pol.PEP. Fig. 110B. 2003_CON_A2 pol.OPT.
(OPT = codon optimized encoding sequence.)

Figures 111A-111D. Fig. 111A. 2003_CON_B
pol.PEP. Fig. 111B. 2003_CON_B pol.OPT. Fig.
20   111C. 2003_B.anc pol.PEP. Fig. 111D. 2003_B.anc
pol.OPT. (OPT = codon optimized encoding sequence.)

Figures 112A-112D. Fig. 112A. 2003_CON_C
pol.PEP. Fig. 112B. 2003_CON_C pol.OPT.

Fig. 112C. 2003_C.anc pol.PEP. Fig. 112D. 2003_C.anc pol.OPT. (OPT = codon optimized encoding sequence.)

5   Figures 113A and 113B. Fig. 113A. 2003_CON_D pol.PEP. Fig. 113B. 2003_CON_D pol.OPT. (OPT = codon optimized encoding sequence.)

Figures 114A and 114B. Fig. 114A. 2003_CON_F1 pol.PEP. Fig. 114B. 2003_CON_F1 pol.OPT. (OPT = codon optimized encoding sequence.)

10   Figures 115A and 115B. Fig. 115A. 2003_CON_F2 pol.PEP. Fig. 115B. 2003_CON_F2 pol.OPT. (OPT = codon optimized encoding sequence.)

Figures 116A and 116B. Fig. 116A. 2003_CON_G pol.PEP. Fig. 116B. 2003_CON_G pol.OPT. 15   (OPT = codon optimized encoding sequence.)

Figures 117A and 117B. Fig. 117A. 2003_CON_H pol.PEP. Fig. 117B. 2003_CON_H pol.OPT. (OPT = codon optimized encoding sequence.)

Figures 118A and 118B. Fig. 118A. 20   2003_CON_01_AE pol.PEP. Fig. 118B. 2003_CON_01_AE pol.OPT. (OPT = codon optimized encoding sequence.)

Figures 119A and 119B. Fig. 119A.
2003_CON_02_AG pol.PEP. Fig. 119B. 2003_CON_02_AG
pol.OPT. (OPT = codon optimized encoding sequence.)

Figures 120A and 120B. Fig. 120A.
5   2003_CON_03_AB pol.PEP. Fig. 120B. 2003_CON_03_AB
pol.OPT. (OPT = codon optimized encoding sequence.)

Figures 121A and 121B. Fig. 121A.
2003_CON_04_CPX pol.PEP. Fig. 121B.
2003_CON_04_CPX pol.OPT. (OPT = codon optimized
10   encoding sequence.)

Figures 122A and 122B. Fig. 122A.
2003_CON_06_CPX pol.PEP. Fig. 122B.
2003_CON_06_CPX pol.OPT. (OPT = codon optimized
encoding sequence.)

15   Figures 123A and 123B. Fig. 123A.
2003_CON_08_BC pol.PEP. Fig. 123B. 2003_CON_08_BC
pol.OPT. (OPT = codon optimized encoding sequence.)

Figures 124A and 124B. Fig. 124A.
2003_CON_10_CD pol.PEP. Fig. 124B. 2003_CON_10_CD
20   pol.OPT. (OPT = codon optimized encoding sequence.)

Figures 125A and 125B. Fig. 125A.
2003_CON_11_CPX pol.PEP. Fig. 125B.

2003_CON_11_CPX pol.OPT.   (OPT = codon optimized
encoding sequence.)

      Figures 126A and 126B.   Fig. 126A.
   2003_CON_12_BF pol.PEP.   Fig. 126B.   2003_CON_12_BF
5   pol.OPT.   (OPT = codon optimized encoding sequence.)

      Figures 127A and 127B.   Fig. 127A.
   2003_CON_14_BG pol.PEP.   Fig. 127B.   2003_CON_14_BG
   pol.OPT.   (OPT = codon optimized encoding sequence.)

               DETAILED DESCRIPTION OF THE INVENTION

10   The present invention relates to an immunogen
   that induces antibodies that neutralize a wide
   spectrum of human immunodeficiency virus (HIV)
   primary isolates and/or that induces a T cell
   response.   The immunogen comprises at least one
15   consensus or ancestral immunogen (e.g., Env, Gag,
   Nef or Pol), or portion or variant thereof.   The
   invention also relates to nucleic acid sequences
   encoding the consensus or ancestral immunogen, or
   portion or variant thereof.   The invention further
20   relates to methods of using both the immunogen and
   the encoding sequences.   While the invention is
   described in detail with reference to specific
   consensus and ancestral immunogens (for example, to
   a group M consensus Env), it will be appreciated
25   that the approach described herein can be used to
   generate a variety of consensus or ancestral

immunogens (for example, envelopes for other HIV-1
groups (e.g., N and O)).

   In accordance with one embodiment of the
invention, a consensus *env* gene can be constructed
5   by generating consensus sequences of *env* genes for
each subtype of a particular HIV-1 group (group M
being classified into subtypes A-D, F-H, J an K),
for example, from sequences in the Los Alamos HIV
Sequence Database (using, for example, MASE
10   (Multiple Aligned Sequence Editor)). A consensus
sequence of all subtype consensuses can then be
generated to avoid heavily sequenced subtypes
(Gaschen et al, Science 296:2354-2360 (2002), Korber
et al, Science 288:1789-1796 (2000)). In the case
15   of the group M consensus *env* gene described in
Example 1 (designated CON6), five highly variable
regions from a CRF08_BC recombinant strain (98CN006)
(V1, V2, V4, V5 and a region in cytoplasmic domain
of gp41) are used to fill in the missing regions in
20   the sequence (see, however, corresponding regions
for Con-S). For high levels of expression, the
codons of consensus or ancestral genes can be
optimized based on codon usage for highly expressed
human genes (Haas et al, Curr. Biol. 6:315-324
25   (2000), Andre et al, J. Virol. 72:1497-1503 (1998)).
   With the Year 1999 consensus group M *env* gene,
CON6, it has been possible to demonstrate induction
of superior T cell responses by CON6 versus wild-
type B and C *env* by the number of ELISPOT
30   γ-interferon spleen spot forming cells and the

38

number of epitopes recognized in two strains of mice (Tables 1 and 2 show the data in BALB/c mice). The ability of CON6 Env protein to induce neutralizing antibodies to HIV-1 primary isolates has been

5   compared to that of several subtype B Env. The target of neutralizing antibodies induced by CON6 includes several non-B HIV-1 strains.

Table 1. T cell epitope mapping of CON6, JRFL and 96ZM651
Env immunogen in BALB/c mice

Peptide	Immunogen			T cell response
	CON6	JRFL (B)	96ZM651 (C)	
CON 6 (group M consensus)				
16 DTEVHNVWATHACVP	+		+	CD4
48 KNSSEYYRLINCNTS 49 EYYRLINCNTSAITQ	+		+	CD4
53 CPKVSFEPIPIHYCA 54 SFEPIPIHYCAPAGF	+			CD4
62 NVSTVQCTHGIKPVV	+			CD4
104 ETITLPCRIKQIINM 105 LPCRIKQIINMWQGV	+			CD8
130 GIVQQQSNLLRAIEA 131 VQQSNLLRAIEAQQHL	+			CD4
134 AQQHLLQLTVWGIKQLQ 135 LQLTVWGIKQLQARVL	+			CD4
Subtype B (MN)				
6223 AKAYDTEVHNVWATQ 6224 DTEVHNVWATQACVP	+			CD4
6261 ACPKISFEPIPIHYC 6262 ISFEPIPIHYCAPAG	+			CD4
6286 RKRIHIGPGRAFYTT 6287 HIGPGRAFYTTKNII		+		CD8
6346 IVQQQNNLLRAIEAQ 6347 QNNLLRAIEAQQHML	+			CD4
Subtype C (Chn19)				
4834 VPVWKEAKTTLFCASDAKSY			+	CD4
4836 GKEVHNVWATHACVPTDPNP	+		+	CD4
4848 SSENSSEYYRLINCNTSAIT	+		+	CD4
4854 STVQCTHGIKPVVSTQLLLN	+			CD4
4884 QQSNLLRAIEAQQHLLQLTV	+			CD4
4885 AQQHLLQLTVWGIKQLQTRV	+			CD4

Table 2.   T cell epitope mapping of CON6.gp120
immunogen in C57BL/6 mice

Peptide	Peptide sequence	T cell response
CON 6 (consensus)		
2	GIQRNCQHLWRWGTM	CD8
3	NCQHLWRWGTMILGM	
16	DTEVHNVWATHACVP	CD4
53	CPKVSFEPIPIHYCA	CD4
97	FYCNTSGLFNSTWMF	CD8
99	FNSTWMFNGTYMFNG	CD8
Subtype B (MN)		
6210	GIRRNYQHWWGWGTM	CD8
6211	NYQHWWGWGTMLLGL	
6232	NMWKNNMVEQMHEDI	CD4
6262	ISFEPIPIHYCAPAG	CD4
6290	NIIGTIRQAHCNISR	CD4
6291	TIRQAHCNISRAKWN	
Subtype C (Chn 19)		
4830	MRVTGIRKNYQHLWRWGTML	CD8
5446	RWGTMLLGMLMICSAAEN	CD8
4836	GKEVHNVWATHACVPTDPNP	CD4
4862	GDIRQAHCNISKDKWNETLQ	CD4
4888	LLGIWGCSGKLICTTTVPWN	CD8

For the Year 2000 consensus group M *env* gene,

5 Con-S, the Con-S envelope has been shown to be as

immunogenic as the CON6 envelope gene in T cell γ

interferon ELISPOT assays in two strains of mice

(the data for C57BL/6 are shown in Fig. 27).
Furthermore, in comparing CON6 and Con-S gp140 Envs
as protein immunogens for antibody in guinea pigs
(Table 3), both gp140 Envs were found to induce

5   antibodies that neutralized subtype B primary
isolates. However, Con-S gp140 also induced robust
neutralization of the subtype C isolates TV-1 and DU
123 as well as one subtype A HIV-1 primary isolate,
while CON6 did not.

TABLE 3   Ability of Group M Consensus CON6 and Con-S Envs to Induce Neutralization of HIV-1 Primary Isolates

HIV-1 Isolate (Subtype)	CON6 gp140CF				CON6 gp140 CFI				CONS gp140 CFI			
	\<Guinea Pig Number\>											
	770	771	772	775	781	783	784	786	776	777	778	780
BX08(B)	520	257	428	189	218	164	>540	199	>540	>540	>540	>540
QH0692 (B)	46	55	58	77	<20	91	100	76	109	<20	<20	<20
SS1196(B)	398	306	284	222	431	242	>540	351	>540	296	>540	>540
JRLFL(B)	<20	<20	<20	<20	<20	169	<20	<20	<20	<20	<20	<20
BG1168(B)	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20
3988(B)	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20
6101(B)	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20
TV-1(C)	<20	<20	<20	<20	<20	<20	<20	<20	356	439	>540	>54
DU123(C)	<20	<20	71	74	<20	72	<20	<20	176	329	387	378
DU172(C)	<20	<20	96	64	<20	<20	<20	<20	<20	235	<20	213
ZM18108.6(C)	ND	ND	ND	ND	<20	<20	<20	<20	84	61	86	43

ZM14654.7(C)	ND	ND	ND	<20	<20	<20	<20	<20	<20	<20	30	<20
DU151(C)	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20
DU422(C)	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20
DU156(C)	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20
92RWO20(A)	<20	<20	<20	<20	44	<20	<20	116	204	95	177
92UG037(A)	<20	30	<20	<20	<20	<20	<20	<20	<20	<20	<2

‡ 50% Neutralization titers after 4th or 5th immunizations

Year 2000 Con-S 140CFI.ENV sequence is shown in Fig. 26A. Gp140 CFI refers to an HIV-1 envelope design in which the cleavage-site is deleted (C), the fusion-site is deleted (F) and the gp41 immunodominant region is deleted (I), in addition to the deletion of transmembrane and cytoplasmic domains. The codon-optimized Year 2000 Con-S 140 CFI sequence is shown in Fig. 26B.

As the next iteration of consensus immunogens, and in recognition of the fact that a practical HIV-1 immunogen can be a polyvalent mixture of either

5 several subtype consensus genes, a mixture of subtype and consensus genes, or a mixture of centralized genes and wild type genes, a series of 11 subtype consensus, and wild type genes have been designed from subtypes A, B, C, CRF AE01, and G as

10 well as a group M consensus gene from Year 2003 Los Alamos National Database sequences. The wild type sequences were chosen either because they were known to come from early transmitted HIV-1 strains (those strains most likely to be necessary to be protected

15 against by a vaccine) or because they were the most recently submitted strains in the database of that subtype. These nucleotide and amino acid sequences are shown in Figures 28-38 (for all 140CF designs shown, 140CF gene can be flanked with the 5'

20 sequence "TTCAGTCGACGGCCACC" that contains a Kozak sequence (GCCACCATGG/A) and *Sal*I site and 3' sequence of TAAAGATCTTACAA containing stop codon and *Bgl*II site). Shown in Figures 39-62 are 2003 centralized (consensus and ancestral) HIV-1 envelope

25 proteins and the codon optimized gene sequences.

Major differences between CON6 gp140 (which does not neutralize non-clade B HIV strains) and Con-S gp140 (which does induce antibodies that neutralize non-clade B HIV strains) are in Con-S V1,

30 V2, V4 and V5 regions. For clade B strains, peptides of the V3 region can induce neutralizing

antibodies (Haynes et al, J. Immunol. 151:1646-1653

(1993)). Thus, construction of Th-V1, Th-V2, Th-

V4, Th-V5 peptides can be expected to give rise to

the desired broadly reactive anti-non-clade B

5   neutralizing antibodies. Therefore, the Th-V

peptides set forth in Table 4 are contemplated for

use as a peptide immunogen(s) derived from Con-S

gp140. The gag Th determinant (GTH, Table 4) or any

homologous GTH sequence in other HIV strains, can be

10   used to promote immunogenicity and the C4 region of

HIV gp120 can be used as well (KQIINMWQVVGKAMYA) or

any homologous C4 sequence from other HIV strains

(Haynes et al, J. Immunol. 151:1646-1653 (1993)).

Con-S V1, V2, V4, V5 peptides with an N-terminal

15   helper determinant can be used singly or together,

when formulated in a suitable adjuvant such as

Corixa's RC529 (Baldridge et al, J. Endotoxin Res.

8:453-458 (2002)), to induce broadly cross reactive

neutralizing antibodies to non-clade B isolates.

20

Table 4		
1)	GTH Con-S V1 132-150	YKRWIILGLNKIVRMYTNVNVTNTTNNTEEKGEIKN
2)	GTH Con-S V2 157-189	YKRWIILGLNKIVRMYTEIRDKKQKVYALFYRLDVVPIDDNNNNSSNYR
3)	GTH Con-S V3 294-315	YKRWIILGLNKIVRMYTRPNNNTRKSIRIGPGQAFYAT
4)	GTH Con-S V4 381-408	YKRWIILGLNKIVRMYNTSGLFNSTWIGNGTKNNNNTNDTITLP
5)	GTH Con-S V5 447-466	YKRWIILGLNKIVRMYRDGGNNNTNETEIFRPGGGD
6)	GTH Con-6 V1 132-150	YKRWIILGLNKIVRMYNVRNVSSNGTETDNEEIKN
7)	GTH Con-6 V2 157-196	YKRWIILGLNKIVRMYTELRDKKQKVYALFYRLDVVPIDDKNSSEISGKNSSEYYR
8)	GTH-Con6 V3 301-322	YKRWIILGLNKIVRMYTRPNNNTRKSIHIGPGQAFYAT
9)	GTH Con-6 V4 388-418	YKRWIILGLNKIVRMYNTSGLFNSTWMFNGTYMFNGTKDNSETITLP
10	GTH Con 6 V5 457-477	YKRWIILGLNKIVRMYRDGGNNSNKNKTETFRPGGGD

It will be appreciated that the invention
includes portions and variants of the sequences
specifically disclosed herein. For example, forms
of codon optimized consensus encoding sequences can
be constructed as gp140CF, gp140 CFI, gp120 or gp160
forms with either gp120/41 cleaved or uncleaved.
For example, and as regards the consensus and
ancestral envelope sequences, the invention
encompasses envelope sequences devoid of V3.
Alternatively, V3 sequences can be selected from
preferred sequences, for example, those described in
U.S. Application No. 10/431,596 and U.S. Provisional
Application No. 60/471,327. In addition, an optimal
immunogen for breadth of response can include
mixtures of group M consensus *gag*, *pol*, *nef* and *env*
encoding sequences, and as well as consist of

mixtures of subtype consensus or ancestral encoding
sequences for *gag*, *pol*, *nef* and *env* HIV genes. For
dealing with regional differences in virus strains,
an efficacious mixture can include mixtures of
5   consensus/ancestral and wild type encoding
sequences.

.   A consensus or ancestral envelope of the
invention can be been "activated" to expose
intermediate conformations of neutralization
10   epitopes that normally are only transiently or less
well exposed on the surface of the HIV virion. The
immunogen can be a "frozen" triggered form of a
consensus or ancestral envelope that makes available
specific epitopes for presentation to B lymphocytes.
15   The result of this epitope presentation is the
production of antibodies that broadly neutralize
HIV.   (Attention is directed to WO 02/024149 and to
the activated/triggered envelopes described
therein.)

20   The concept of a fusion intermediate immunogen
is consistent with observations that the gp41 HR-2
region peptide, DP178, can capture an uncoiled
conformation of gp41 (Furata et al, Nature Struct.
Biol. 5:276 (1998)), and that formalin-fixed HIV-
25   infected cells can generate broadly neutralizing
antibodies (LaCasse et al, Science 283:357 (1997)).
Recently a monoclonal antibody against the coiled-
coil region bound to a conformational determinant of
gp41 in HR1 and HR2 regions of the coiled-coil gp41
30   structure, but did not neutralize HIV (Jiang et al,
J. Virol. 10213 (1998)). However, this latter study

proved that the coiled-coil region is available for antibody to bind if the correct antibody is generated.

5   The immunogen of one aspect of the invention comprises a consensus or ancestral envelope either in soluble form or anchored, for example, in cell vesicles or in liposomes containing translipid bilayer envelope. To make a more native envelope, gp140 or gp160 consensus or ancestral sequences can

10   be configured in lipid bilayers for native trimeric envelope formation. Alternatively, triggered gp160 in aldrithio 1-2 inactivated HIV-1 virions can be used as an immunogen. The gp160 can also exist as a recombinant protein either as gp160 or gp140 (gp140

15   is gp160 with the transmembrane region and possibly other gp41 regions deleted). Bound to gp160 or gp140 can be recombinant CCR5 or CXCR4 co-receptor proteins (or their extracellular domain peptide or protein fragments) or antibodies or other ligands

20   that bind to the CXCR4 or CCR5 binding site on gp120, and/or soluble CD4, or antibodies or other ligands that mimic the binding actions of CD4. Alternatively, vesicles or liposomes containing CD4, CCR5 (or CXCR4), or soluble CD4 and peptides

25   reflective of CCR5 or CXCR4 gp120 binding sites. Alternatively, an optimal CCR5 peptide ligand can be a peptide from the N-terminus of CCR5 wherein specific tyrosines are sulfated (Bormier et al, Proc. Natl. Acad. Sci. USA 97:5762 (2001)). The

30   triggered immunogen may not need to be bound to a membrane but may exist and be triggered in solution.

Alternatively, soluble CD4 (sCD4) can be replaced by
an envelope (gp140 or gp160) triggered by CD4
peptide mimetopes (Vitra et al, Proc. Natl. Acad.
Sci. USA 96:1301 (1999)). Other HIV co-receptor
5   molecules that "trigger" the gp160 or gp140 to
undergo changes associated with a structure of gp160
that induces cell fusion can also be used. Ligation
of soluble HIV gp140 primary isolate HIV 89.6
envelope with soluble CD4 (sCD4) induced
10 conformational changes in gp41.

    In one embodiment, the invention relates to an
immunogen that has the characteristics of a receptor
(CD4)-ligated consensus or ancestral envelope with
CCR5 binding region exposed but unlike CD4-ligated
15 proteins that have the CD4 binding site blocked,
this immunogen has the CD4 binding site exposed
(open). Moreover, this immunogen can be devoid of
host CD4, which avoids the production of potentially
harmful anti-CD4 antibodies upon administration to a
20 host.

    The immunogen can comprise consensus or
ancestral envelope ligated with a ligand that binds
to a site on gp120 recognized by an A32 monoclonal
antibodies (mab) (Wyatt et al, J. Virol. 69:5723
25 (1995), Boots et al, AIDS Res. Hum. Retro. 13:1549
(1997), Moore et al, J. Virol. 68:8350 (1994),
Sullivan et al, J. Virol. 72:4694 (1998), Fouts et
al, J. Virol. 71:2779 (1997), Ye et al, J. Virol.
74:11955 (2000)). One A32 mab has been shown to
30 mimic CD4 and when bound to gp120, upregulates
(exposes) the CCR5 binding site (Wyatt et al, J.

Virol. 69:5723 (1995)). Ligation of gp120 with such
a ligand also upregulates the CD4 binding site and
does not block CD4 binding to gp120.
Advantageously, such ligands also upregulate the HR-
5   2 binding site of gp41 bound to cleaved gp120,
uncleaved gp140 and cleaved gp41, thereby further
exposing HR-2 binding sites on these proteins - each
of which are potential targets for anti-HIV
neutralizing antibodies.

10   In a specific aspect of this embodiment, the
immunogen comprises soluble HIV consensus or
ancestral gp120 envelope ligated with either an
intact A32 mab, a Fab2 fragment of an A32 mab, or a
Fab fragment of an A32 mab, with the result that the
15   CD4 binding site, the CCR5 binding site and the HR-2
binding site on the consensus or ancestral envelope
are exposed/upregulated. The immunogen can comprise
consensus or ancestral envelope with an A32 mab (or
fragment thereof) bound or can comprise consensus or
20   ancestral envelope with an A32 mab (or fragment
thereof) bound and cross-linked with a cross-linker
such as .3% formaldehyde or a heterobifunctional
cross-linker such as DTSSP (Pierce Chemical
Company). The immunogen can also comprise uncleaved
25   consensus or ancestral gp140 or a mixture of
uncleaved gp140, cleaved gp41 and cleaved gp120. An
A32 mab (or fragment thereof) bound to consensus or
ancestral gp140 and/or gp120 or to gp120 non-
covalently bound to gp41, results in upregulation
30   (exposure) of HR-2 binding sites in gp41, gp120 and
uncleaved gp140. Binding of an A32 mab (or fragment

thereof) to gp120 or gp140 also results in
upregulation of the CD4 binding site and the CCR5
binding site. As with gp120 containing complexes,
complexes comprising uncleaved gp140 and an A32 mab
5   (or fragment thereof) can be used as an immunogen
uncross-linked or cross-linked with cross-linker
such as .3% formaldehyde or DTSSP. In one
embodiment, the invention relates to an immunogen
comprising soluble uncleaved consensus or ancestral
10 gp140 bound and cross linked to a Fab fragment or
whole A32 mab, optionally bound and cross-linked to
an HR-2 binding protein.

The consensus or ancestral envelope protein
triggered with a ligand that binds to the A32 mab
15 binding site on gp120 can be administered in
combination with at least a second immunogen
comprising a second envelope, triggered by a ligand
that binds to a site distinct from the A32 mab
binding site, such as the CCR5 binding site
20 recognized by mab 17b. The 17b mab (Kwong et al,
Nature 393:648 (1998) available from the AIDS
Reference Repository, NIAID, NIH) augments sCD4
binding to gp120. This second immunogen (which can
also be used alone or in combination with triggered
25 immunogens other than that described above) can, for
example, comprise soluble HIV consensus or ancestral
envelope ligated with either the whole 17b mab, a
Fab2 fragment of the 17b mab, or a Fab fragment of
the 17b mab. It will be appreciated that other CCR5
30 ligands, including other antibodies (or fragments
thereof), that result in the CD4 binding site being

exposed can be used in lieu of the 17b mab. This
further immunogen can comprise gp120 with the 17b
mab, or fragment thereof, (or other CCR5 ligand as
indicated above) bound or can comprise gp120 with
5   the 17b mab, or fragment thereof, (or other CCR5
ligand as indicated above) bound and cross-linked
with an agent such as .3% formaldehyde or a
heterobifunctional cross-linker, such as DTSSP
(Pierce Chemical Company). Alternatively, this
10 further immunogen can comprise uncleaved gp140
present alone or in a mixture of cleaved gp41 and
cleaved gp120. Mab 17b, or fragment thereof (or
other CCR5 ligand as indicated above) bound to gp140
and/or gp120 in such a mixture results in exposure
15 of the CD4 binding region. The 17b mab, or fragment
thereof, (or other CCR5 ligand as indicated above)
gp140 complexes can be present uncross-linked or
cross-linked with an agent such as .3% formaldehyde
or DTSSP.

20   Soluble HR-2 peptides, such as T649Q26L and
DP178, can be added to the above-described complexes
to stabilize epitopes on consensus gp120 and gp41 as
well as uncleaved consensus gp140 molecules, and can
be administered either cross-linked or uncross-
25 linked with the complex.

A series of monoclonal antibodies (mabs) have
been made that neutralize many HIV primary isolates,
including, in addition to the 17b mab described
above, mab IgG1b12 that binds to the CD4 binding
30 site on gp120(Roben et al, J. Virol. 68:482 (1994),
Mo et al, J. Virol. 71:6869 (1997)), mab 2G12 that

binds to a conformational determinant on gp120
(Trkola et al, J. Virol. 70:1100 (1996)), and mab
2F5 that binds to a membrane proximal region of gp41
(Muster et al, J. Virol. 68:4031 (1994)).

5   As indicated above, various approaches can be
used to "freeze" fusogenic epitopes in accordance
with the invention. For example, "freezing" can be
effected by addition of the DP-178 or T-649Q26L
peptides that represent portions of the coiled coil
10   region, and that when added to CD4-triggered
consensus or ancestral envelope, result in
prevention of fusion (Rimsky et al, J. Virol.
72:986-993 (1998)). HR-2 peptide bound consensus or
ancestral gp120, gp140, gp41 or gp160 can be used as
15   an immunogen or crosslinked by a reagent such as
DTSSP or DSP (Pierce Co.), formaldehyde or other
crosslinking agent that has a similar effect.

       "Freezing" can also be effected by the addition
of 0.1% to 3% formaldehyde or paraformaldehyde, both
20   protein cross-linking agents, to the complex, to
stabilize the CD4, CCR5 or CXCR4, HR-2 peptide gp160
complex, or to stabilize the "triggered" gp41
molecule, or both (LaCasse et al, Science 283:357-
362 (1999)).

25   Further, "freezing" of consensus or ancestral
gp41 or gp120 fusion intermediates can be effected
by addition of heterobifunctional agents such as DSP
(dithiobis[succimidylproprionate]) (Pierce Co.
Rockford, ILL., No. 22585ZZ) or the water soluble
30   DTSSP (Pierce Co.) that use two NHS esters that are
reactive with amino groups to cross link and

stabilize the CD4, CCR5 or CXCR4, HR-2 peptide gp160
complex, or to stabilize the "triggered" gp41
molecule, or both.

Analysis of T cell immune responses in
5   immunized or vaccinated animals and humans shows
that the envelope protein is normally not a main
target for T cell immune response although it is the
only gene that induces neutralizing antibodies.
HIV-1 Gag, Pol and Nef proteins induce a potent T
10   cell immune response. Accordingly, the invention
includes a repertoire of consensus or ancestral
immunogens that can induce both humoral and cellular
immune responses. Subunits of consensus or
ancestral sequences can be used as T or B cell
15   immunogens. (See Examples 6 and 7, and Figures
referenced therein, and Figures 63-127.

The immunogen of the invention can be
formulated with a pharmaceutically acceptable
carrier and/or adjuvant (such as alum) using
20   techniques well known in the art. Suitable routes
of administration of the present immunogen include
systemic (e.g. intramuscular or subcutaneous).
Alternative routes can be used when an immune
response is sought in a mucosal immune system (e.g.,
25   intranasal).

The immunogens of the invention can be
chemically synthesized and purified using methods
which are well known to the ordinarily skilled
artisan. The immunogens can also be synthesized by
30   well-known recombinant DNA techniques. Nucleic
acids encoding the immunogens of the invention can

be used as components of, for example, a DNA vaccine
wherein the encoding sequence is administered as
naked DNA or, for example, a minigene encoding the
immunogen can be present in a viral vector. The
5   encoding sequence can be present, for example, in a
replicating or non-replicating adenoviral vector, an
adeno-associated virus vector, an attenuated
mycobacterium tuberculosis vector, a Bacillus
Calmette Guerin (BCG) vector, a vaccinia or Modified
10   Vaccinia Ankara (MVA) vector, another pox virus
vector, recombinant polio and other enteric virus
vector, Salmonella species bacterial vector,
Shigella species bacterial vector, Venezuelean
Equine Encephalitis Virus (VEE) vector, a Semliki
15   Forest Virus vector, or a Tobacco Mosaic Virus
vector. The encoding sequence, can also be
expressed as a DNA plasmid with, for example, an
active promoter such as a CMV promoter. Other live
vectors can also be used to express the sequences of
20   the invention. Expression of the immunogen of the
invention can be induced in a patient's own cells,
by introduction into those cells of nucleic acids
that encode the immunogen, preferably using codons
and promoters that optimize expression in human
25   cells. Examples of methods of making and using DNA
vaccines are disclosed in U.S. Pat. Nos. 5,580,859,
5,589,466, and 5,703,055.

         The composition of the invention comprises an
immunologically effective amount of the immunogen of
30   this invention, or nucleic acid sequence encoding
same, in a pharmaceutically acceptable delivery

system. The compositions can be used for prevention
and/or treatment of immunodeficiency virus
infection. The compositions of the invention can be
formulated using adjuvants, emulsifiers,

5   pharmaceutically-acceptable carriers or other
ingredients routinely provided in vaccine
compositions. Optimum formulations can be readily
designed by one of ordinary skill in the art and can
include formulations for immediate release and/or

10   for sustained release, and for induction of systemic
immunity and/or induction of localized mucosal
immunity (e.g, the formulation can be designed for
intranasal administration). The present
compositions can be administered by any convenient

15   route including subcutaneous, intranasal, oral,
intramuscular, or other parenteral or enteral route.
The immunogens can be administered as a single dose
or multiple doses. Optimum immunization schedules
can be readily determined by the ordinarily skilled

20   artisan and can vary with the patient, the
composition and the effect sought.

        The invention contemplates the direct use of
both the immunogen of the invention and/or nucleic
acids encoding same and/or the immunogen expressed

25   as minigenes in the vectors indicated above. For
example, a minigene encoding the immunogen can be
used as a prime and/or boost.

        The invention includes any and all amino acid
sequences disclosed herein and, where applicable, CF

30   and CFI forms thereof, as well as nucleic acid

sequences encoding same (and nucleic acids
complementary to such encoding sequences).

    Certain aspects of the invention can be
described in greater detail in the non-limiting
5   Examples that follows.

                        EXAMPLE 1

        Artificial HIV-1 Group M Consensus Envelope

EXPERIMENTAL DETAILS

10   *Expression of CON6 gp120 and gp140 proteins in
    recombinant vaccinia viruses (VV).* To express and
    purify the secreted form of HIV-1 CON6 envelope
    proteins, CON6 gp120 and gp140CF plasmids were
    constructed by introducing stop codons after the
15 gp120 cleavage site (REKR) and before the
    transmembrane domain (YIKIFIMIVGGLIGLRIVFAVLSIVN),
    respectively. The gp120/gp41 cleavage site and
    fusion domain of gp41 were deleted in the gp140CF
    protein. Both CON6 gp120 and gp140CF DNA constructs
20 were cloned into the pSC65 vector (from Bernard
    Moss, NIH, Bethesda, MD) at SalI and KpnI
    restriction enzyme sites. This vector contains the
    lacZ gene that is controlled by the p7.5 promoter.
    A back-to-back P E/L promoter was used to express
25 CON6 *env* genes. BSC-1 cells were seeded at 2×10^5
    in each well in a 6-well plate, infected with wild-
    type vaccinia virus (WR) at a MOI of 0.1 pfu/cell,
    and 2 hr after infection, pSC65-derived plasmids

containing CON6 *env* genes were transfected into the
VV-infected cells and recombinant (r) VV selected as
described (Moss and Earl, Current Protocols in
Molecular Biology, eds, Ausubel et al (John Wiley &
5   Sons, Inc. Indianapolis, IN) pp. 16.15.1-16.19.9
(1998)). Recombinant VV that contained the CON6 *env*
genes were confirmed by PCR and sequencing analysis.
Expression of the CON6 envelope proteins was
confirmed by SDS-PAGE and Western blot assay.
10   Recombinant CON6 gp120 and gp140CF were purified
with agarose *galanthus Nivalis* lectin beads (Vector
Labs, Burlingame, CA), and stored at -70°C until use.
Recombinant VV expressing JRFL (vCB-28) or 96ZM651
(vT241R) gp160 were obtained from the NIH AIDS
15   Research and Reference Reagent Program (Bethesda,
MD).

     *Monoclonal Antibodies and gp120 Wild-type
     Envelopes.* Human mabs against a conformational
20   determinant on gp120 (A32), the gp120 V3 loop (F39F)
and the CCR5 binding site (17b) were the gifts of
James Robinson (Tulane Medical School, New Orleans,
LA) (Wyatt et al, Nature 393;705-711 (1998), Wyatt
et al, J. Virol. 69:5723-5733 (1995)). Mabs 2F5,
25   447, b12, 2G12 and soluable CD4 were obtained from
the NIH AIDS Research and Reference Reagent Program
(Bethesda, MD) (Gorny et al, J. Immunol. 159:5114-
5122 (1997), Nyambi et al, J. Virol. 70:6235-6243
(1996), Purtscher et al, AIDS Res. Hum. Retroviruses
30   10:1651-1658 (1994), Trkola et al, J. Virol 70:1100-
1108 (1996)). T8 is a murine mab that maps to the

gp120 C1 region (a gift from P. Earl, NIH, Bethesda,
MD). BaL (subtype B), 96ZM651 (subtype C), and
93TH975 (subtype E) gp120s were provided by QBI,
Inc. and the Division of AIDS, NIH. CHO cell lines

5   that express 92U037 (subtype A) and 93BR029 (subtype
F) gp140 (secreted and uncleaved) were obtained from
NICBS, England.

     Surface Plasmon Resonance Biosensor (SPR)

10   *Measurements and ELISA.* SPR biosensor measurements
were determined on a BIAcore 3000 instrument
(BIAcore Inc., Uppsala, Sweden) instrument and data
analysis was performed using BIAevaluation 3.0
software (BIAcore Inc, Upsaala, Sweden). Anti-gp120

15   mabs (T8, A32, 17b, 2G12) or sCD4 in 10mM Na-acetate
buffer, pH 4.5 were directly immobilized to a CM5
sensor chip using a standard amine coupling protocol
for protein immobilization. FPLC purified CON6
gp120 monomer or gp140CF oligomer recombinant

20   proteins were flowed over CM5 sensor chips at
concentrations of 100 and 300 μg/ml, respectively.
A blank in-line reference surface (activated and de-
activated for amine coupling) or non-bonding mab
controls were used to subtract non-specific or bulk

25   responses. Soluble 89.6 gp120 and irrelevant IgG
was used as a positive and negative control
respectively and to ensure activity of each mab
surface prior to injecting the CON6 Env proteins.
Binding of CON6 envelope proteins was monitored in

30   real-time at 25°C with a continuous flow of PBS (150
mM NaCl, 0.005% surfactant P20), pH 7.4 at 10-30

μl/min. Bound proteins were removed and the sensor
surfaces were regenerated following each cycle of
binding by single or duplicate 5-10 μl pulses of
regeneration solution (10 mM glycine-HCl, pH 2.9).

5   ELISA was performed to determine the reactivity of
various mabs to CON6 gp120 and gp140CF proteins as
described (Haynes et al, AIDS Res. Hum. Retroviruses
11:211-221 (1995)). For assay of human mab binding
to rgp120 or gp140 proteins, end-point titers were

10   defined as the highest titer of mab (beginning at 20
μg/ml) at which the mab bound CON6 gp120 and gp140CF
Env proteins $\geq$ 3 fold over background control (non-
binding human mab).

15   *Infectivity and coreceptor usage assays.* HIV-
1/SG3Δenv and CON6 or control *env* plasmids were
cotransfected into human 293T cells. Pseudotyped
viruses were harvested, filtered and p24
concentration was quantitated (DuPont/NEN Life

20   Sciences, Boston, MA). Equal amounts of p24 (5 ng)
for each pseudovirion were used to infect JC53-BL
cells to determine the infectivity (Derdeyn e al, J.
Virol. 74:8358-8367 (2000), Wei et al, Antimicrob
Agents Chemother. 46:1896-1905 (2002)). JC53-BL

25   cells express CD4, CCR5 and CXCR4 receptors and
contain a β-galactosidase (β-gal) gene stably
integrated under the transcriptional control of an
HIV-1 long terminal repeat (LTR). These cells can
be used to quantify the infectious titers of

30   pseudovirion stocks by staining for β-gal expression

and counting the number of blue cells (infectious
units) per microgram of p24 of pseudovirons (IU/μg
p24) (Derdeyn e al, J. Virol. 74:8358-8367 (2000),
Wei et al, Antimicrob Agents Chemother. 46:1896-1905
5   (2002)). To determine the coreceptor usage of the
CON6 *env* gene, JC53BL cells were treated with 1.2 μM
AMD3100 and 4 μM TAK-799 for 1 hr at 37°C then
infected with equal amounts of p24 (5 ng) of each
Env pseudotyped virus. The blockage efficiency was
10   expressed as the percentage of the infectious units
from blockage experiments compared to that from
control culture without blocking agents. The
infectivity from control group (no blocking agent)
was arbitrarily set as 100%.
15

     Immunizations. All animals were housed in the
Duke University Animal Facility under AALAC
guidelines with animal use protocols approved by the
Duke University Animal Use and Care Committee.
20   Recombinant CON6 gp120 and gp140CF glycoproteins
were formulated in a stable emulsion with RIBI-CWS
adjuvant based on the protocol provided by the
manufacturer (Sigma Chemical Co., St. Louis, MO).
For induction of anti-envelope antibodies, each of
25   four out-bred guinea pigs (Harlan Sprague, Inc.,
Chicago, IL) was given 100 µg either purified CON6
gp120 or gp140CF subcutaneously every 3 weeks (total
of 5 immunizations). Serum samples were heat-
inactivated (56°C, 1 hr), and stored at -20°C until
30   use.

For induction of anti-envelope T cell
responses, 6-8 wk old female BALB/c mice (Frederick
Cancer Research and Developmental Center, NCI,
Frederick, MD) were immunized i.m. in the quadriceps
5   with 50 μg plasmid DNA three times at a 3-week
interval. Three weeks after the last DNA
immunization, mice were boosted with 10^7 PFU of rVV
expressing Env proteins. Two weeks after the boost,
all mice were euthanized and spleens were removed
10   for isolation of splenocytes.

      Neutralization assays. Neutralization assays
were performed using either a MT-2 assay as
described in Bures et al, AIDS Res. Hum.
15   Retroviruses 16:2019-2035 (2000), a luciferase-based
multiple replication cycle HIV-1 infectivity assay
in 5.25.GFP.Luc.M7 cells using a panel of HIV-1
primary isolates (Bures et al, AIDS Res. Hum.
Retroviruses 16:2019-2035 (2000), Bures et al, J.
20   Virol. 76:2233-2244 (2002)), or a syncytium (fusion
from without) inhibition assay using inactivated
HIV-1 virions (Rossio et al, J. Virol. 72:7992-8001
(1998)). In the luciferase-based assay,
neutralizing antibodies were measured as a function
25   of a reduction in luciferase acitivity in
5.25.EGFP.Luc.M7 cells provided by Nathaniel R.
Landau, Salk Institute, La Jolla, CA (Brandt et al,
J. Biol. Chem. 277:17291-17299 (2002)). Five
hundred tissue culture infectious dose 50 ($TCID_{50}$) of
30   cell-free virus was incubated with indicated serum

dilutions in 150 μl (1 hr, at 37°C) in triplicate in
96-well flat-bottom culture plates. The
5.25.EGFP.Luc.M7 cells were suspended at a density
of 5 x 10^5/ml in media containing DEAE dextran (10

5   μg/ml). Cells (100 μl) were added and until 10% of
cells in control wells (no test serum sample) were
positive for GFP expression by fluorescence
microscopy. At this time the cells were
concentrated 2-fold by removing one-half volume of

10 media. A 50 μl suspension of cells was transferred
to 96-well white solid plates (Costar, Cambridge,
MA) for measurement of luciferase activity using
Bright-Glo™ substrate (Promega, Madison, WI) on a
Wallac 1420 Multilabel Counter (PerkinElmer Life

15 Sciences, Boston, MA). Neutralization titers in the
MT-2 and luciferase assays were those where ≥ 50%
virus infection was inhibited. Only values that
titered beyond 1:20 (i.e. >1:30) were considered
significantly positive. The syncytium inhibition

20 "fusion from without" assay utilized HIV-1
aldrithiol-2 (AT-2) inactivated virions from HIV-1
subtype B strains ADA and AD8 (the gift of Larry
Arthur and Jeffrey Lifson, Frederick Research Cancer
Facility, Frederick, MD) added to SupT1 cells, with

25 syncytium inhibition titers determined as those
titers where ≥90% of syncytia were inhibited
compared to prebleed sera.

    Enzyme linked immune spot (ELISPOT) assay.
30 Single-cell suspensions of splenocytes from

individual immunized mice were prepared by mincing
and forcing through a 70 μm Nylon cell strainer (BD
Labware, Franklin Lakes, NJ). Overlapping Env
peptides of CON6 gp140 (159 peptides, 15mers
5   overlapping by 11) were purchased from Boston
Bioscence, Inc (Royal Oak, MI). Overlapping Env
peptides of MN gp140 (subtype B; 170 peptides,
15mers overlapping by 11) and Chn19 gp140 (subtype
C; 69 peptides, 20mers overlapping by 10) were
10   obtained from the NIH AIDS Research and Reference
Reagent Program (Bethesda, MD). Splenocytes (5
mice/group) from each mouse were stimulated *in vitro*
with overlapping Env peptides pools from CON6,
subtype B and subtype C Env proteins. 96-well PVDF
15   plates (MultiScreen-IP, Millipore, Billerica, MA)
were coated with anti-IFN-γ mab (5 μg/ml, AN18;
Mabtech, Stockholm, Sweden). After the plates were
blocked at 37°C for 2 hr using complete Hepes
buffered RPMI medium, 50μl of the pooled overlapping
20   envelope peptides (13 CON6 and MN pools, 13-14
peptides in each pool; 9 Chn19 pool, 7-8 peptide in
each pool) at a final concentration of 5 μg/ml of
each were added to the plate. Then 50 μl of
splenocytes at a concentration of 1.0×10^7/ml were
25   added to the wells in duplicate and incubated for 16
hr at 37°C with 5% CO_2. The plates were incubated
with 100 μl of a 1:1000 dilution of streptavidin
alkaline phosphatase (Mabtech, Stockholm, Sweden),
and purple spots developed using 100 μl of BCIP/NBT
30   (Plus) Alkaline Phosphatase Substrate (Moss,

Pasadena, MD). Spot forming cells (SFC) were measured using an Immunospot counting system (CTL Analyzers, Cleveland, OH). Total responses for each envelope peptide pool are expressed as SFCs per 10^6

5   splenocytes.

RESULTS

       CON6 Envelope Gene Design, Construction and

10   *Expression.* An artificial group M consensus *env* gene (CON6) was constructed by generating consensus sequences of *env* genes for each HIV-1 subtype from sequences in the Los Alamos HIV Sequence Database, and then generating a consensus sequence of all

15   subtype consensuses to avoid heavily sequenced subtypes (Gaschen et al, Science 296:2354-2360 (2002), Korber et al, Science 288:1789-1796 (2000)). Five highly variable regions from a CRF08_BC recombinant strain (98CN006) (V1, V2, V4, V5 and a

20   region in cytoplasmic domain of gp41) were then used to fill in the missing regions in CON6 sequence. The CON6 V3 region is group M consensus (Figure 1A). For high levels of expression, the codons of CON6 *env* gene were optimized based on codon usage for

25   highly expressed human genes (Haas et al, Curr. Biol. 6:315-324 (2000), Andre et al, J. Virol. 72:1497-1503 (1998)). (See Fig. 1D.) The codon optimized CON6 env gene was constructed and subcloned into pcDNA3.1 DNA at EcoR I and BamH I

30   sites (Gao et al, AIDS Res. Hum. Retroviruses, 19:817-823 (2003)). High levels of protein

expression were confirmed with Western-blot assays
after transfection into 293T cells. To obtain
recombinant CON6 Env proteins for characterization
and use as immunogens, rVV was generated to express
5   secreted gp120 and uncleaved gp140CF (Figure 1B).
Purity for each protein was ≥90% as determined by
Coomassie blue gels under reducing conditions
(Figure 1C).

10   *CD4 Binding Domain and Other Wild-type HIV-1*
Epitopes are Preserved on CON6 Proteins. To
determine if CON6 proteins can bind to CD4 and
express other wild-type HIV-1 epitopes, the ability
of CON6 gp120 and gp140CF to bind soluble(s) CD4, to
15   bind several well-characterized anti-gp120 mabs, and
to undergo CD4-induced conformational changes was
assayed. First, BIAcore CM5 sensor chips were
coated with either sCD4 or mabs to monitor their
binding activity to CON6 Env proteins. It was found
20   that both monomeric CON6 gp120 and oligomeric
gp140CF efficiently bound sCD4 and anti-gp120 mabs
T8, 2G12 and A32, but did not constitutively bind
mab 17b, that recognizes a CD4 inducible epitope in
the CCR5 binding site of gp120 (Figures 2A and 2B).
25   Both sCD4 and A32 can expose the 17b binding epitope
after binding to wild-type gp120 (Wyatt et al,
Nature 393;705-711 (1998), Wyatt et al, J. Virol.
69:5723-5733 (1995)). To determine if the 17b
epitope could be induced on CON6 Envs by either sCD4
30   or A32, sCD4, A32 and T8 were coated on sensor
chips, then CON6 gp120 or gp140CF captured, and mab

17b binding activity monitored. After binding sCD4
or mab A32, both CON6 gp120 and gp140CF were
triggered to undergo conformational changes and
bound mab 17b (Figures 2C and 2D). In contrast,

5   after binding mab T8, the 17b epitope was not
exposed (Figures 2C and 2D). ELISA was next used to
determine the reactivity of a panel of human mabs
against the gp120 V3 loop (447, F39F), the CD4
binding site (b12), and the gp41 neutralizing

10   determinant (2F5) to CON6 gp120 and gp140CF (Figure
2E). Both CON6 rgp120 and rgp140CF proteins bound
well to neutralizing V3 mabs 447 and F39F and to the
potent neutralizing CD4 binding site mab b12. Mab
2F5, that neutralizes HIV-1 primary isolates by

15   binding to a C-terminal gp41 epitope, also bound
well to CON6 gp140CF (Figure 2E).

         *CON6 env Gene is Biologically Functional and
Uses CCR5 as its Coreceptor.* To determine whether

20   CON6 envelope gene is biologically functional, it
was co-transfected with the *env*-defective SG3
proviral clone into 293T cells. The pseudotyped
viruses were harvested and JC53BL cells infected.
Blue cells were detected in JC53-BL cells infected

25   with the CON6 Env pseudovirions, suggesting that
CON6 Env protein is biologically functional (Figure
3A). However, the infectious titers were 1-2 logs
lower than that of pseudovirions with either YU2 or
NL4-3 wild-type HIV-1 envelopes.

30   The co-receptor usage for the CON6 *env* gene was
next determined. When treated with CXCR4 blocking

agent AMD3100, the infectivity of NL4-3 Env-
pseudovirons was blocked while the infectivity of
YU2 or CON6 Env-pseudovirons was not inhibited
(Figure 3B). In contrast, when treated with CCR5

5   blocking agent TAK-779, the infectivity of NL4-3
Env-pseudovirons was not affected, while the
infectivity of YU2 or CON6 Env-pseudovirons was
inhibited. When treated with both blocking agents,
the infectivity of all pseudovirions was inhibited.

10   Taken together, these data show that the CON6
envelope uses the CCR5 co-receptor for its entry
into target cells.

     Reaction of CON6 gp120 With Different Subtype

15   *Sera.* To determine if multiple subtype linear
epitopes are preserved on CON6 gp120, a recombinant
Env protein panel (gp120 and gp140) was generated.
Equal amounts of each Env protein (100 ng) were
loaded on SDS-polyacrylamide gels, transferred to

20   nitrocellulose, and reacted with subtype A through G
patient sera as well as anti-CON6 gp120 guinea pig
sera (1:1,000 dilution) in Western blot assays. For
each HIV-1 subtype, four to six patient sera were
tested. One serum representative for each subtype

25   is shown in Figure 4.
          It was found that whereas all subtype sera
tested showed variable reactivities among Envs in
the panel, all group M subtype patient sera reacted
equally well with CON6 gp120 Env protein,

30   demonstrating that wild-type HIV-1 Env epitopes
recognized by patient sera were well preserved on

the CON6 Env protein. A test was next made as to
whether CON6 gp120 antiserum raised in guinea pigs
could react to different subtype Env proteins. It
was found that the CON6 serum reacted to its own and
5   other subtype Env proteins equally well, with the
exception of subtype A Env protein (Figure 4).

*Induction of T Cell Responses to CON6, Subtype
B and Subtype C Envelope Overlapping Peptides.* To
10,   compare T cell immune responses induced by CON6 Env
immunogens with those induced by subtype specific
immunogens, two additional groups of mice were
immunized with subtype B or subtype C DNAs and with
corresponding rVV expressing subtype B or C envelope
15   proteins. Mice immunized with subtype B (JRFL) or
subtype C (96ZM651) Env immunogen had primarily
subtype-specific T cell immune responses (Figure 5).
IFN-γ SFCs from mice immunized with JRFL (subtype B)
immunogen were detected after stimulation with
20   subtype B (MN) peptide pools, but not with either
subtype C (Chn19) or CON6 peptide pools. IFN-γ SFCs
from mice immunized with 96ZM651 (subtype C)
immunogen were detected after the stimulation with
both subtype C (Chn19) and CON6 peptide pools, but
25   not with subtype B (MN) peptide pools. In contrast,
IFN-γ SFCs were identified from mice immunized with
CON6 Env immunogens when stimulated with either CON6
peptide pools as well as by subtype B or C peptide
pools (Figure 5). The T cell immune responses
30   induced by CON6 gp140 appeared more robust than

those induced by CON6 gp120. Taken together, these
data demonstrated that CON6 gp120 and gp140CF
immunogens were capable of inducing T cell responses
that recognized T cell epitopes of wild-type subtype
5   B and C envelopes.

      Induction of Antibodies by Recombinant CON6
gp120 and gp140CF Envelopes that Neutralize HIV-1
Subtype B and C Primary Isolates. To determine if
10   the CON6 envelope immunogens can induce antibodies
that neutralize HIV-1 primary isolates, guinea pigs
were immunized with either CON6 gp120 or gp140CF
protein. Sera collected after 4 or 5 immunizations
were used for neutralization assays and compared to
15   the corresponding prebleed sera. Two AT-2
inactivated HIV-1 isolates (ADA and AD8) were tested
in syncytium inhibition assays (Table 5A). Two
subtype B SHIV isolates, eight subtype B primary
isolates, four subtype C, and one each subtype A, D,
20   and E primary isolates were tested in either the MT-
2 or the luciferase-based assay (Table 5B). In the
syncytium inhibition assay, it was found that
antibodies induced by both CON 6 gp120 and gp140CF
proteins strongly inhibited AT-2 inactivated ADA and
25   AD8-induced syncytia (Table 5A). In the MT-2 assay,
weak neutralization of 1 of 2 SHIV isolates (SHIV
SF162P3) by two gp120 and one gp140CF sera was found
(Table 5B). In the luciferase-based assay, strong
neutralization of 4 of 8 subtype B primary isolates
30   (BXO8, SF162, SS1196, and BAL) by all gp120 and
gp140CF sera was found, and weak neutralization of 2

of 8 subtype B isolates (6101, 0692) by most gp120
and gp140CF sera was found.   No neutralization was
detected against HIV-1 PAVO (Table 5B).   Next, the
CON6 anti-gp120 and gp140CF sera were tested against

5   four subtype C HIV-1 isolates, and weak
neutralization of 3 of 4 isolates (DU179, DU368, and
S080) was found, primarily by anti-CON6 gp120 sera.
One gp140CF serum, no. 653, strongly neutralized
DU179 and weakly neutralized S080 (Table 5B).

10 Finally, anti-CON6 Env sera strongly neutralized a
subtype D isolate (93ZR001), weakly neutralized a
subtype E (CM244) isolate, and did not neutralize a
subtype A (92RW020) isolate.

Table 5A

Ability of HIV-1 Group M Consensus Envelope CON6 Proteins to Induce Fusion Inhibiting Antibodies

Guinea Pig No.	Immunogen	Syncytium Inhibition antibody titer[1]	
		AD8	ADA
646	gp120	270	270
647	gp120	90	90
648	gp120	90	270
649	gp120	90	90
Geometric Mean Titer		119	156
650	gp140	270	270
651	gp140	90	90
652	gp140	≥810	810
653	gp140	270	90
Geometric Mean Titer		270	207

15

[1]Reciprocal serum dilution at which HIV-induced syncytia of Sup T1 cells was
inhibited by >90% compared to pre-immune serum. All prebleed sera were negative
(titer <10).

Table 5B

Ability of Group M Consensus HIV-1 Envelope CON6 gp120 and gp140CF Proteins to Induce Antibodies that Neutralize HIV Primary Isolates

HIV Isolate (Subtype)	CON6 gp120 Protein Guinea Pig No.					CON6 gp140CF Protein Guinea Pig No.					Controls		
	646	647	648	649	GMT	650	651	652	653	GMT	TriMab2‡	CD4-IgG2	HIV+ Serum
SHIV 89.6P*(B)	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	NT	NT	NT
SHIV SF162P3*(B)	<20	30	48	<20	<20	27	<20	<20	<20	<20	NT	0.2µg/ml	NT
BX08(B)	270	183	254	55	102	199	64	229	150	187	0.7µg/ml	NT	2384#
6101(B)	<20	38	35	<20	<20	<20	90	72	73	39	1.1µg/ml	NT	NT
BG1168(B)	<20	<20	<20	<20	<20	40	<20	<20	25	<20	2.7µg/ml	NT	NT
0692(B)	31	32	34	<20	24	28	33	30	45	33	0.8µg/ml	NT	769
PAVO(B)	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	2.9µg/ml	NT	NT
SF162(B)	2,146	308	110	282	379	206	5,502	15,098	174	1,313	NT	NT	>540
SS1196(B)	206	26	148	59	83	381	401	333	81	253	NT	NT	301#
BAL(B)	123	90	107	138	113	107	146	136	85	116	NT	NT	3307
92RW020(A)	<20	<20	<20	<20	<20	<20	<20	<20	<20	<20	NT	NT	693
DU179(C)	<20	43	<20	24	<20	<20	<20	24	515	33	NT	0.8µg/ml	NT
DU368(C)	25	35	62	<20	27	<20	<20	<20	23	<20	NT	2.3µg/ml	NT
S021(C)	<20	<20	33	<20	<20	<20	<20	<20	<20	<20	NT	8.3µg/ml	NT
S080(C)	24	37	70	41	40	<20	<20	<20	52	<20	NT	3.4µg/ml	NT
93ZR001(D)	275	144	126	114	154	306	195	129	173	191	NT	NT	693
CM244(E)	35	43	64	ND	46	31	25	27	25	26	NT	NT	693

*MT-2 Assay; All other HIV isolates were tested in the M7-luciferase assay.

HIV-1 isolates QH0692, SS1196, SF162, 6101, BX08, BG1168, BAL were assayed with post-injection 5 serum; other HIV-1 isolates were assayed with post-injection 4 serum. ND = not done.

HIV+ sera was either HIV-1+ human serum (LEH3) or an anti-gp120 guinea pig serum (#) with known neutralizing activity for HIV-1 isolate SS1196. GMT = geometric mean titer of four animals per group. Neutralizing titers reported are after subtraction of any background neutralization in prebleed sera.

‡TriMab2 = a mixture of human mabs 2F5, b12, 2G12.

CONCLUSIONS

        The production of an artificial HIV-1 Group M
consensus env genes (encoding sequences) (CON6 and
5   Con-S) have been described that encodes a functional
Env protein that is capable of utilizing the CCR5
co-receptor for mediating viral entry. Importantly,
these Group M consensus envelope genes could induce
T and B cell responses that recognized epitopes of
10 subtype B and C HIV-1 primary isolates. In
addition, Con-S induces antibodies that strongly
neutralize Subtype-C and A HIV-1 strains (see
Table 3).

        The correlates of protection to HIV-1 are not
15 conclusively known. Considerable data from animal
models and studies in HIV-1-infected patients
suggest the goal of HIV-1 vaccine development should
be the induction of broadly-reactive CD4+ and CD8+
anti-HIV-1 T cell responses (Letvin et al, Annu.
20 Rev. Immunol. 20:73-99 (2002)) and high levels of
antibodies that neutralize HIV-1 primary isolates of
multiple subtypes (Mascola et al, J. Virol. 73:4009-
4018 (1999), Mascola et al, Nat. Med. 6:270-210
(2000)).

25   The high level of genetic variability of HIV-1
has made it difficult to design immunogens capable
of inducing immune responses of sufficient breadth
to be clinically useful. Epitope based vaccines for
T and B cell responses (McMichael et al, Vaccine
30 20:1918-1921 (2002), Sbai et al, Curr. Drug Targets
Infect, Disord. 1:303-313 (2001), Haynes, Lancet

348:933-937 (1996)), constrained envelopes
reflective of fusion intermediates (Fouts et al,
Proc. Natl. Acad. Sci. USA 99:11842-22847 (2002)),
as well as exposure of conserved high-order

5   structures for induction of anti-HIV-1 neutralizing
antibodies have been proposed to overcome HIV-1
variability (Roben et al, J. Virol. 68:4821-4828
(1994), Saphire et al, Science 293:1155-1159
(2001)). However, with the ever-increasing

10   diversity and rapid evolution of HIV-1, the virus is
a rapidly moving complex target, and the extent of
complexity of HIV-1 variation makes all of these
approaches problematic. The current most common
approach to HIV-1 immunogen design is to choose a

15   wild-type field HIV-1 isolate that may or may not be
from the region in which the vaccine is to be
tested. Polyvalent envelope immunogens have been
designed incorporating multiple envelope immunogens
(Bartlett et al, AIDS 12:1291-1300 (1998), Cho et

20   al, J. Virol. 75:2224-2234 (2001)).

The above-described study tests a new strategy
for HIV-1 immunogen design by generating a group M
consensus *env* gene (CON6) with decreased genetic
distance between this candidate immunogen and wild-

25   type field virus strains. The CON6 *env* gene was
generated for all subtypes by choosing the most
common amino acids at most positions (Gaschen et al,
Science 296:2354-2360 (2002), Korber et al, Science
288:1789-1796 (2000)). Since only the most common

30   amino acids were used, the majority of antibody and
T cell epitopes were well preserved. Importantly,

the genetic distances between the group M consensus
env sequence and any subtype *env* sequences was about
15%, which is only half of that between wild-type
subtypes (30%) (Gaschen et al, Science 296:2354-2360

5   (2002)). This distance is approximately the same as
that among viruses within the same subtype.
Further, the group M consensus *env* gene was also
about 15% divergent from any recombinant viral *env*
gene, as well, since CRFs do not increase the

10   overall genetic divergence among subtypes.
        Infectivity of CON6-Env pseudovirions was
confirmed using a single-round infection system,
although the infectivity was compromised, indicating
the artificial envelope was not in an "optimal"

15   functional conformation, but yet was able to mediate
virus entry. That the CON6 envelope used CCR5 (R5)
as its coreceptor is important, since majority of
HIV-1 infected patients are initially infected with
R5 viruses.

20   BIAcore analysis showed that both CON6 gp120
and gp140CF bound sCD4 and a number of mabs that
bind to wild-type HIV-1 Env proteins. The
expression of the CON6 gp120 and 140CF proteins that
are similar antigenically to wild-type HIV-1

25   envelopes is an important step in HIV-1 immunogen
development. However, many wild-type envelope
proteins express the epitopes to which potent
neutralizing human mabs bind, yet when used as
immunogens themselves, do not induce broadly

30   neutralizing anti-HIV-1 antibodies of the
specificity of the neutralizing human mabs.

The neutralizing antibody studies were encouraging in that both CON6 gp120, CON6 gp140CF and Con-S gp140CFI induced antibodies that neutralized select subtype B, C and D HIV-1 primary
5   isolates, with Con-S gp140CFI inducing the most robust neutralization of non-subtype B primary HIV isolates. However, it is clear that the most difficult-to-neutralize primary isolates (PAVO, 6101, BG1168, 92RW020, CM244) were either only
10   weakly or not neutralized by anti-CON6 gp120 or gp140 sera (Table 4b). Nonetheless, the Con-S envelope immunogenicity for induction of neutralizing antibodies is promising, given the breadth of responses generated with the Con-S
15   subunit gp140CFI envelope protein for non-subtype B HIV isolates. Previous studies with poxvirus constructs expressing gp120 and gp160 have not generated high levels of neutralizing antibodies (Evans et al, J. Infect. Dis. 180:290-298 (1999),
20   Polacino et al, J. Virol. 73:618-630 (1999), Ourmanov et al, J. Virol. 74:2960-2965 (2000), Pal et al, J. Virol 76:292-302 (2002), Excler and Plotkin, AIDS 11(Suppl A):S127-137 (1997). rVV expressing secreted CON6 gp120 and gp140 have been
25   constructed and antibodies that neutralize HIV-1 primary isolates induced. An HIV neutralizing antibody immunogen can be a combination of Con-S gp140CFI, or subunit thereof, with immunogens that neutralize most subtype B isolates.

The structure of an oligomeric gp140 protein is
critical when evaluating protein immunogenicity. In
this regard, study of purified CON6 gp140CF proteins
5   by fast performance liquid chromatography (FPLC) and
analytical ultracentrifiguration has demonstrated
that the purified gp140 peak consists predominantly
of trimers with a small component of dimers.

Thus, centralized envelopes such as CON6, Con-S
10   or 2003 group M or subtype consensus or ancestral
encoding sequences described herein, are attractive
candidates for preparation of various potentially
"enhanced" envelope immunogens including CD4-Env
complexes, constrained envelope structures, and
15 · trimeric oligomeric forms. The ability of CON6-
induced T and B cell responses to protect against
HIV-1 infection and/or disease in SHIV challenge
models will be studied in non-human primates.

The above study has demonstrated that
20   artificial centralized HIV-1 genes such as group M
consensus *env* gene (CON6) and Con-S can also induce
T cell responses to T cell epitopes in wild-type
subtype B and C Env proteins as well as to those on
group M consensus Env proteins (Figure 5). While
25   the DNA prime and rVV boost regimen with CON6
gp140CF immunogen clearly induced IFN-γ producing T
cells that recognized subtype B and C epitopes,
further studies are needed to determine if
centralized sequences such as are found in the CON6
30   envelope are significantly better at inducing cross-

clade T cell responses than wild-type HIV-1 genes
(Ferrari et al, Proc. Natl. Acad. Sci. USA 94:1396-
1401 (1997), Ferrari et al, AIDS Res. Hum.
Retroviruses 16:1433-1443 (2000)). However, the
5   fact that CON6 (and Con-S env encoding sequence)
prime and boosted splenocyte T cells recognized HIV-
1 subtype B and C T cell epitopes is an important
step in demonstration that CON6 (and Con-S) can
induce T cell responses that might be clinically
10   useful.

Three computer models (consensus, ancestor and
center of the tree (COT)) have been proposed to
generate centralized HIV-1 genes (Gaschen et al,
Science 296:2354-2360 (2002), Gao et al, Science
15   299:1517-1518 (2003), Nickle et al, Science
299:1515-1517 (2003), Korber et al, Science
288:1789-1796 (2000). They all tend to locate at
the roots of the star-like phylogenetic trees for
most HIV-1 sequences within or between subtypes. As
20   experimental vaccines, they all can reduce the
genetic distances between immunogens and field virus
strains. However, consensus, ancestral and COT
sequences each have advantages and disadvantages
(Gaschen et al, Science 296:2354-2360 (2002), Gao et
25   al, Science 299:1517-1518 (2003), Nickle et al,
Science 299:1515-1517 (2003). Consensus and COT
represent the sequences or epitopes in sampled
current wild-type viruses and are less affected by
outliers HIV-1 sequences, while ancestor represents
30   ancestral sequences that can be significantly
affected by outlier sequences. However, at present,

it is not known which centralized sequence can serve
as the best immunogen to elicit broad immune
responses against diverse HIV-1 strains, and studies
are in progress to test these different strategies.

5   Taken together, the data have shown that the
HIV-1 artificial CON6 and Con-S envelope can induce
T cell responses to wild-type HIV-1 epitopes, and
can induce antibodies that neutralize HIV-1 primary
isolates, thus demonstrating the feasibility and
10 promise of using artificial centralized HIV-1
sequences in HIV-1 vaccine design.

<div align="center">EXAMPLE 2</div>

<div align="center">HIV-1 Subtype C Ancestral and Consensus Envelope</div>
15 <div align="center">Glycoproteins</div>

EXPERIMENTAL DETAILS
     HIV-1 subtype C ancestral and consensus *env*
genes were obtained from the Los Alamos HIV
Molecular Immunology Database (http://hiv-
20 web.lanl.gov/immunology), codon-usage optimized for
mammalian cell expression, and synthesized (Fig. 6).
To ensure optimal expression, a Kozak sequence
(GCCGCCGCC) was inserted immediately upstream of the
initiation codon. In addition to the full-length
25 genes, two truncated *env* genes were generated by
introducing stop codons immediately after the gp41
membrane-spanning domain (IVNR) and the gp120/gp41
cleavage site (REKR), generating gp140 and gp120
form of the glycoproteins, respectively (Fig. 8).

Genes were tested for integrity in an *in vitro* transcription/translation system and expressed in mammalian cells. To determine if the ancestral and consensus subtype C envelopes were capable of

5   mediating fusion and entry, *gp160* and *gp140* genes were co-transfected with an HIV-1/SG3Δ*env* provirus and the resulting pseudovirions tested for infectivity using the JC53-BL cell assay (Fig. 7). Co-receptor usage and envelope neutralization

10   sensitivity were also determined with slight modifications of the JC53-BL assay. Codon-usage optimized and rev-dependent 96ZAM651 *env* genes were used as contemporary subtype C controls.

RESULTS

15

Codon-optimized subtype C ancestral and consensus envelope genes (*gp160, gp140, gp120*) express high levels of *env* glycoprotein in mammalian cells (Fig. 9).

20   Codon-optimized subtype C *gp160* and *gp140* glycoproteins are efficiently incorporated into virus particles. Western Blot analysis of sucrose-purified pseudovirions reveals ten-fold higher levels of virion incorporation of the codon-

25   optimized envelopes compared to that of a rev-dependent contemporary envelope controls (Fig. 10A).

Virions pseudotyped with either the subtype C consensus *gp160* or *gp140* envelope were more infectious than pseudovirions containing the

30   corresponding gp160 and gp140 ancestral envelopes.

Additionally, *gp160* envelopes were consistently more infectious than their respective *gp140* counterparts (Fig. 10B).

5   Both subtype C ancestral and consensus envelopes utilize CCR5 as a co-receptor to mediate virus entry (Fig. 11).

The infectivity of subtype C ancestral and consensus *gp160* containing pseudovirions was neutralized by plasma from subtype C infected
10   patients. This suggests that these artificial envelopes possess a structure that is similar to that of native HIV-1 *env* glycoproteins and that common neutralization epitopes are conserved. No significant differences in neutralization potential
15   were noted between subtype C ancestral and consensus *env* glycoproteins (gp160) (Fig. 12).

CONCLUSIONS

HIV-1 subtype C viruses are among the most prevalent circulating isolates, representing
20   approximately fifty percent of new infections worldwide. Genetic diversity among globally circulating HIV-1 strains poses a challenge for vaccine design. Although HIV-1 Env protein is highly variable, it can induce both humoral and cellular
25   immune responses in the infected host. By analyzing 70 HIV-1 complete subtype C *env* sequences, consensus and ancestral subtype C *env* genes have been generated. Both sequences are roughly equidistant from contemporary subtype C strains and thus

expected to induce better cross-protective immunity.
A reconstructed ancestral or consensus sequence
derived-immunogen minimizes the extent of genetic
differences between the vaccine candidate and
5   contemporary isolates. However, consensus and
ancestral subtype C *env* genes differ by 5% amino
acid sequences. Both consensus and ancestral
sequences have been synthesized for analyses.
Codon-optimized subtype C ancestral and consensus
10   envelope genes have been constructed and the *in
vitro* biological properties of the expressed
glycoproteins determined. Synthetic subtype C
consensus and ancestral *env* genes express
glycoproteins that are similar in their structure,
15   function and antigenicity to contemporary subtype C
wild-type envelope glycoproteins.

EXAMPLE 3

Codon-Usage Optimization of Consensus of Subtype C
20   *gag* and *nef* Genes (C.con.gag and C.con.nef)

Subtype C viruses have become the most
prevalent viruses among all subtypes of Group M
viruses in the world. More than 50% of HIV-1
25   infected people are currently carrying HIV-1 subtype
C viruses. In addition, there is considerable
intra-subtype C variability: different subtype C
viruses can differ by as much as 10%, 6%, 17% and

82

16% of their Gag, Pol, Env and Nef proteins,
respectively. Most importantly, the subtype C
viruses from one country can vary as much as the
viruses isolated from other parts of the world. The
5   only exceptions are HIV-1 strains from India/China,
Brazil and Ethiopia/Djibouti where subtype C appears
to have been introduced more recently. Due to the
high genetic variability of subtype C viruses even
within a single country, an immunogen based on a
10   single virus isolate may not elicit protective
immunity against other isolates circulating in the
same area.

Thus *gag* and *nef* gene sequences of subtype C
viruses were gathered to generate consensus
15   sequences for both genes by using a 50% consensus
threshold. To avoid a potential bias toward founder
viruses, only one sequence was used from
India/China, Brazil and Ethiopia/Djibouti,
respectively, to generate the subtype C consensus
20   sequences (C.con.gag and C.con.nef). The codons of
both C.con.gag and C.con.nef genes were optimized
based on the codon usage of highly expressed human
genes. The protein expression following transfection
into 293T cells is shown in Figure 13. As can be
25   seen, both consensus subtype C Gag and Nef proteins
were expressed efficiently and recognized by Gag-
and Nef-specific antibodies. The protein expression
levels of both C.con.gag and C.con.nef genes are
comparible to that of native subtype *env* gene
30   (96ZM651).

<u>EXAMPLE 4</u>

Synthesis of a Full Length "*Consensus of the Consensus env* Gene with Consensus Variable Regions" (CON-S)

In the synthesized "consensus of the consensus" *env* gene (CON6), the variable regions were replaced with the corresponding regions from a contemporary subtype C virus (98CN006). A further con/con gene has been designed that also has consensus variable regions (CON-s). The codons of the Con-S *env* gene were optimized based on the codon usage of highly expressed human genes. (See Figs. 14A and 14B for amino acid sequences and nucleic acid sequences, respectfully.)

Paired oligonucleotides (80-mers) which overlap by 20 bp at their 3' ends and contain invariant sequences at their 5' and 3' ends, including the restriction enzyme sites EcoRI and BbsI as well as BsmBI and BamHI, respectively, were designed. BbsI and BamHI are Type II restriction enzymes that cleave outside of their recognition sequences. They have been positioned in the oligomers in such a way that they cleave the first four resides adjacent to the 18 bp invariant region, leaving 4 base 5' overhangs at the end of each fragment for the following ligation step. 26 paired oligomers were linked individually using PCR and primers complimentary to the 18 bp invariant sequences.

Each pair was cloned into pGEM-T (Promega) using the

T/A cloning method and sequenced to confirm the

absence of inadvertent mutations/deletions. pGEM-T

subclones containing the proper inserts were then

5   digested, run on a 1% agarose gel, and gel purified

(Qiagen). Four individual 108-mers were ligated

into pcDNA3.1 (Invitrogen) in a multi-fragment

ligation reaction. The four-way ligations occurred

among groups of fragments in a stepwise manner from

10   the 5' to the 3' end of the gene. This process was

repeated until the entire gene was reconstructed in

the pcDNA3.1 vector.

A complete Con-S gene was constructed by

ligating the codon usage optimized oligo pairs

15   together. To confirm its open reading frame, an *in

vitro* transcription and translation assay was

performed. Protein products were labeled by S^{35}-

methionine during the translation step, separated on

a 10% SDS-PAGE, and detected by radioautography.

20   Expected size of the expressed Con-S gp160 was

identified in 4 out of 7 clones (Fig. 14C).

CONs Env protein expression in the mammalian

cells after transfected into 293T cells using a

Western blot assay (Figure 15). The expression level

25   of Con-S Env protein is very similar to what was

observed from the previous CON6 *env* clone that

contains the consensus conservative regions and

variable loops from 98CN006 virus isolate.

The Env-pseudovirons was produced by

30   cotransfecting Con-S *env* clone and *env*-deficient SG3

proviral clone into 293T cells. Two days after
transfection, the pseudovirions were harvested and
infected into JC53BL-13 cells. The infectious units
(IU) were determined by counting the blue cells
5   after staining with X-gal in three independent
experiments. When compared with CON6 *env* clone, Con-
S *env* clones produce similar number of IU in JC53BL-
13 cells (Figure 16). The IU titers for both are
about 3 log higher than the SG3 backbone clone
10   control (No Env). However, the titers are also
about 2 log lower than the positive control (the
native HIV-1 *env* gene, NL4-3 or YU2). These data
suggest that both *consensus* group M *env* clones are
biologically functional. Their functionality,
15   however, has been compromised. The functional
consensus *env* genes indicate that these Env proteins
fold correctly, preserve the basic conformation of
the native Env proteins, and are able to be
developed as universal Env immunogens.
20   It was next determined what coreceptor Con-S
Env uses for its entry into JC53-BL cells. When
treated with CXCR4 blocking agent AMD3100, the
infectivity of NL4-3 Env-pseudovirons was blocked
while the infectivity of YU2, Con-S or CON6 Env-
25   pseudovirons was not inhibited. In contrast, when
treated with CCR5 blocking agent TAK779, the
infectivity of NL4-3 Env-pseudovirons was not
affected, while the infectivity of YU2, Con-S or
CON6 Env-pseudovirons was inhibited. When treated
30   with both blocking agents, the infectivity of all
pseudovirions was inhibited. Taken together, these

data show that the Con-S as well as CON6 envelope
uses the CCR5 but not CXCR4 co-receptor for its
entry into target cells.

It was next determined whether CON6 or Con-S
5   Env proteins could be equally efficiently
incorporated in to the pseudovirions. To be able
precisely compare how much Env proteins were
incorporated into the pseudovirions, each
pseudovirions is loaded on SDS-PAGE at the same
10   concentraion: 5μg total protein for cell lysate,
25ng p24 for cell culture supernatant, or 150ng p24
for purified virus stock (concentrated pseudovirions
after super-speed centrifugation). There was no
difference in amounts of Env proteins incorporated
15   in CON6 or Con-S Env-pseudovirions in any
preparations (cell lysate, cell culture supernatant
or purified virus stock) (Figure 17).

<u>EXAMPLE 5</u>

Synthesis of a *Consensus* Subtype A Full Length *env*
20   (A.con.env) Gene

Subtype A viruses are the second most prevalent
HIV-1 in the African continent where over 70% of
HIV-1 infections have been documented. Consensus
25   *gag*, *env* and *nef* genes for subtype C viruses that
are the most prevalent viruses in Africa and in the
world were previously generated. Since genetic
distances between subtype A and C viruses are as
high as 30% in the *env* gene, the cross reactivity or
30   protection between both subtypes will not be

optimal. Two group M consensus *env* genes for all
subtypes were also generated. However, to target
any particular subtype viruses, the subtype specific
consensus genes will be more effective since the
5   genetic distances between subtype consensus genes
and field viruses from the same subtype will be
smaller than that between group M consensus genes
and these same viruses. Therefore, consensus genes
need to be generated for development of subtype A
10   specific immunogens. The codons of the A.con.env
gene were optimized based on the codon usage of
highly expressed human genes. (See Figs. 18A and
18B for amino acid and nucleic acid sequences,
respectively.)

15   Each pair of the oligos has been amplified,
cloned, ligated and sequenced. After the open
reading frame of the A.con *env* gene was confirmed by
an *in vitro* transcription and translation system,
the A.con *env* gene was transfected into the 293T
20   cells and the protein expression and specificity
confirmed with the Western blot assay (Figure 18).
It was then determined whether A.con envelope is
biologically functional. It was co-transfected with
the *env*-defective SG3 proviral clone into 293T
25   cells. The pseudotyped viruses were harvested and
used to infect JC53BL cells. Blue cells were
detected in JC53-BL cells infected with the A.con
Env-pseudovirions, suggesting that A.con Env protein
is biologically functional (Table 6). However, the
30   infectious titer of A.con Env-psuedovirions was
about 7-fold lower than that of pseudovirions with

wild-type subtype C envelope (Table 6). Taken
together, the biological function A.con Env proteins
suggests that it folds correctly and may induce
linear and conformational T and B cell epitopes if
5   used as an Env immunogen.

JC53BL13 (IU/ul)

	3/31/03	4/7/03	4/25/03
	non filtered supt.	0.22μm filtered	0.22μm filtered
A.con +SG3	4	8.5	15.3
96ZM651 +SG3	87	133	104
SG3 backbone	0	0.07	0.03
Neg control	0	0.007	0

Table 6. Infectivity of pseudovirons with A.con env genes

EXAMPLE 6

10   Design of Full Length *"Consensus of the Consensus
gag, pol and nef* Genes" (M.con.gag, M.con.pol and
M.con.nef) and a Subtype C *Consensus* pol Gene
(C.con.pol)

15   For the group M consensus genes, two different
env genes were constructed, one with virus specific
variable regions (CON6) and one with consensus
variable regions (Con-S). However, analysis of T
cell immune responses in immunized or vaccinated
20   animals and humans shows that the *env* gene normally
is not a main target for T cell immune response

although it is the only gene that will induce
neutralizing antibody. Instead, HIV-1 Gag, Pol and
Nef proteins are found to be important for inducing
potent T cell immune responses. To generate a
5   repertoire of immunogens that can induce both
broader humoral and cellular immune responses for
all subtypes, it may be necessary to construct other
group M consensus genes other than env gene alone.
"Consensus of the consensus" gag, pol and nef genes
10   (M.con.gag., M.con.pol and M.con.nef) have been
designed. To generate a subtype consensus pol gene,
the subtype C consensus pol gene (C.con.pol) was
also designed. The codons of the M.con.gag.,
M.con.pol, M.con.nef and C.con.pol. genes were
15   optimized based on the codon usage of highly
expressed human genes. (See Fig. 19 for nucleic
acid and amino acid sequences.)

                         EXAMPLE 7

        Synthetic Subtype B Consensus gag and env Genes

20   EXPERIMENTAL DETAILS
        Subtype B consensus gag and env sequences were
derived from 37 and 137 contemporary HIV-1 strains,
respectively, codon-usage optimized for mammalian
cell expression, and synthesized (Figs. 20A and
25   20B). To ensure optimal expression, a Kozak
sequence (GCCGCCGCC) was inserted immediately
upstream of the initiation codon. In addition to
the full-length env gene, a truncated env gene was
generated by introducing a stop codon immediately

                            90

after the gp41 membrane-spanning domain (IVNR) to
create a *gp145* gene. Genes were tested for
integrity in an *in vitro* transcription/translation
system and expressed in mammalian cells. (Subtype B
5   consensus Gag and Env sequences are set forth in
Figs. 20C and 20D, respectively.)

To determine if the subtype B consensus
envelopes were capable of mediating fusion and
entry, *gp160* and *gp145* genes were co-transfected
10 with an HIV-1/SG3Δenv provirus and the resulting
pseudovirions were tested for infectivity using the
JC53-BL cell assay. JC53-BL cells are a derivative
of HeLa cells that express high levels of CD4 and
the HIV-1 coreceptors CCR5 and CXCR4. They also
15 contain the reporter cassettes of luciferase and β-
galactosidase that are each expressed from an HIV-1
LTR. Expression of the reporter genes is dependent
on production of HIV-1 Tat. Briefly, cells are
seeded into 24-well plates, incubated at 37°C for 24
20 hours and treated with DEAE-Dextran at 37°C for
30min. Virus is serially diluted in 1% DMEM, added
to the cells incubating in DEAE-dextran, and allowed
to incubate for 3 hours at 37°C after which an
additional 500μL of cell media is added to each
25 well. Following a final 48-hour incubation at 37°C,
cells are fixed, stained using X-Gal, and overlaid
with PBS for microscopic counting of blue foci.
Counts for mock-infected wells, used to determine
background, are subtracted from counts for the
30 sample wells. Co-receptor usage and envelope

neutralization sensitivity were also determined with
slight modifications of the JC53-BL assay.

To determine whether the subtype B consensus
Gag protein was capable of producing virus-like

5   particles (VLPs) that incorporated Env
glycoproteins, 293T cells were co-transfected with
subtype B consensus *gag* and *env* genes. 48-hours
post-transfection, cell supernatants containing VLPs
were collected, clarified in a tabletop centrifuge,

10   filtered through a 0.2mM filter, and pellet through
a 20% sucrose cushion. The VLP pellet was
resuspended in PBS and transferred onto a 20-60%
continuous sucrose gradient. Following overnight
centrifugation at 100,000 x g, 0.5 ml fractions were

15   collected and assayed for p24 content. The
refractive index of each fraction was also measured.
Fractions with the correct density for VLPs and
containing the highest levels of p24 were pooled and
pellet a final time. VLP-containing pellets were

20   re-suspended in PBS and loaded on a 4-20% SDS-PAGE
gel. Proteins were transferred to a PVDF membrane
and probed with serum from a subtype B HIV-1
infected individual.

RESULTS

25

Codon-usage optimized, subtype B consensus
envelope (*gp160, gp145*) and *gag* genes express high
levels of glycoprotein in mammalian cells (Fig. 21).
Subtype B gp160 and gp145 glycoproteins are

30   efficiently incorporated into virus particles.

92

Western Blot analysis of sucrose-purified
pseudovirions suggests at least five-fold higher
levels of consensus B envelope incorporation
compared to incorporation of a rev-dependent
5   contemporary envelope (Fig.23A). Virions
pseudotyped with either the subtype B consensus
gp160 or gp145 envelope are more infectious than
pseudovirions containing a rev-dependent
contemporary envelope (Fig. 23 B).
10   Subtype B consensus envelopes utilize CCR5 as
the co-receptor to gain entry into CD4 bearing
target cells (Fig. 22).
        The infectivity of pseudovirions containing the
subtype B consensus gp160 envelope was neutralized
15   by plasma from HIV-1 subtype B infected patients
(Fig. 24C) and neutralizing monoclonal antibodies
(Fig. 24A). This suggests that the subtype B
synthetic consensus B envelopes is similar to native
HIV-1 Env glycoproteins in its overall structure and
20   that common neutralization epitopes remain intact.
Figs. 24B and 24D show neutralization profiles of a
subtype B control envelope (NL4.3 Env).
        Subtype B consensus Gag proteins are able to
bud from the cell membrane and form virus-like
25   particles (Fig. 25A). Co-transfection of the codon-
optimized subtype B consensus gag and gp160 genes
produces VLPs with incorporated envelope (Fig. 25B).

CONCLUSIONS

    The synthetic subtype B consensus *env* and *gag*
genes express viral proteins that are similar in
their structure, function and antigenicity to
5   contemporary subtype B Env and Gag proteins. It is
contemplated that immunogens based on subtype B
consensus genes will elicit CTL and neutralizing
immune responses that are protective against a broad
set of HIV-1 isolates.

10   *   *   *

    All documents and other information sources
cited above are hereby incorporated in their
entirety by reference. Also incorporated by
reference is Liao et al, J. Virol. 78:5270 (2004)).

WHAT IS CLAIMED IS:

1.   An isolated protein comprising the
sequence of amino acids set forth in Fig. 1A.

2.   A nucleic acid comprising a nucleotide
sequence encoding CON6 HIV gp160 protein, wherein
said nucleotide sequence comprises codons optimized
for expression in human cells.

3.   The nucleic acid according to claim 2
wherein said nucleic acid comprises the nucleotide
sequence set forth in Fig. 1D.

4.   A nucleic acid comprising a nucleotide
sequence encoding subtype C ancestral HIV envelope
protein, wherein said nucleotide sequence comprises
codons optimized for expression in human cells.

5.   The nucleic acid according to claim 4
wherein said nucleic acid comprises the nucleotide
sequence set forth in Fig. 6A.

6.   A nucleic acid comprising a nucleotide
sequence encoding subtype C consensus HIV envelope
protein, wherein said nucleotide sequence comprises
codons optimized for expression in human cells.

7.   The nucleic acid according to claim 6
wherein said nucleic acid comprises the nucleotide
sequence set forth in Fig. 6B.

8.   An isolated protein comprising the
sequence of amino acids set forth in Fig. 6C or 6D.

9.   A nucleic acid comprising a nucleotide sequence encoding a subtype C consensus HIV gag protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

10.   The nucleic acid according to claim 9 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 13E.

11.   A nucleic acid comprising a nucleotide sequence encoding a subtype C consensus HIV nef protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

12.   The nucleic acid according to claim 11 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 13F.

13.   A nucleic acid comprising a nucleotide sequence encoding Group M consensus HIV envelope protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

14.   The nucleic acid according to claim 13 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 14B.

15.   A nucleic acid comprising a nucleotide sequence encoding subtype A consensus HIV envelope protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

16.   The nucleic acid according to claim 15
wherein said nucleic acid comprises the nucleotide
sequence set forth in Fig. 18B.

17.   A nucleic acid comprising a nucleotide
sequence encoding Group M consensus HIV gag protein,
wherein said nucleotide sequence comprises codons
optimized for expression in human cells.

18.   The nucleic acid according to claim 17
wherein said nucleic acid comprises the nucleotide
sequence set forth in Fig. 19A .

19.   A nucleic acid comprising a nucleotide
sequence encoding Group M consensus HIV pol protein,
wherein said nucleotide sequence comprises codons
optimized for expression in human cells.

20.   The nucleic acid according to claim 19
wherein said nucleic acid comprises the nucleotide
sequence set forth in Fig. 19B.

21.   A nucleic acid comprising a nucleotide
sequence encoding Group M consensus HIV nef protein,
wherein said nucleotide sequence comprises codons
optimized for expression in human cells.

22.   The nucleic acid according to claim 21
wherein said nucleic acid comprises the nucleotide
sequence set forth in Fig. 19C.

23.   A nucleic acid comprising a nucleotide
sequence encoding subtype C consensus HIV pol

protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

24.   The nucleic acid according to claim 23 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 19D.

25.   A nucleic acid comprising a nucleotide sequence encoding subtype B consensus HIV gag protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

26.   The nucleic acid according to claim 25 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 20A.

27.   A nucleic acid comprising a nucleotide sequence encoding subtype B consensus HIV envelope protein, wherein said nucleotide sequence comprises codons optimized for expression in human cells.

28.   The nucleic acid according to claim 27 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 20B.

29.   An isolated protein comprising the sequence of amino acids set forth in Fig. 20C or 20D.

30.   An isolated protein comprising the sequence of amino acids set forth in Fig. 26A .

31. A nucleic acid comprising a nucleotide sequence that encodes the protein according to claim 30.

32. The nucleic acid according to claim 31 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 26B.

33. An isolated protein comprising the sequence of amino acids set forth in Fig. 28B.

34. A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 33.

35. The nucleic acid sequence according to claim 34 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 28C.

36. An isolated protein comprising the sequence of amino acids set forth in Fig. 29B.

37. A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 36.

38. The nucleic acid sequence according to claim 37 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 29C.

39. An isolated protein comprising the sequence of amino acids set forth in Fig. 30B.

40. A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 39.

41.   The nucleic acid sequence according to claim 40 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 30C.

42.   An isolated protein comprising the sequence of amino acids set forth in Fig. 31B.

43.   A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 42.

44.   The nucleic acid sequence according to claim 43 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 31C.

45.   An isolated protein comprising the sequence of amino acids set forth in Fig. 32B.

46.   A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 45.

47.   The nucleic acid sequence according to claim 46 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 32C.

48.   An isolated protein comprising the sequence of amino acids set forth in Fig. 33B.

49.   A nucleic acid comprising a nucleotide sequence encoding the protein according to claim 48.

50.   The nucleic acid sequence according to claim 49 wherein said nucleic acid comprises the nucleotide sequence set forth in Fig. 33C.

51. An isolated protein comprising the
sequence of amino acids set forth in Fig. 34B.

52. A nucleic acid comprising a nucleotide
sequence encoding the protein according to claim 51.

53. The nucleic acid sequence according to
claim 52 wherein said nucleic acid comprises the
nucleotide sequence set forth in Fig. 34C.

54. An isolated protein comprising the
sequence of amino acids set forth in Fig. 35B.

55. A nucleic acid comprising a nucleotide
sequence encoding the protein according to claim 54.

56. The nucleic acid sequence according to
claim 55 wherein said nucleic acid comprises the
nucleotide sequence set forth in Fig. 35C.

57. An isolated protein comprising the
sequence of amino acids set forth in Fig. 36B.

58. A nucleic acid comprising a nucleotide
sequence encoding the protein according to claim 57.

59. The nucleic acid sequence according to
claim 58 wherein said nucleic acid comprises the
nucleotide sequence set forth in Fig. 36C.

60. An isolated protein comprising the
sequence of amino acids set forth in Fig. 37B.

61. A nucleic acid comprising a nucleotide
sequence encoding the protein according to claim 60.

62.   The nucleic acid sequence according to
claim 61 wherein said nucleic acid comprises the
nucleotide sequence set forth in Fig. 37C.

63.   An isolated protein comprising the
sequence of amino acids set forth in Fig. 38B.

64.   A nucleic acid comprising a nucleotide
sequence encoding the protein according to claim 63.

65.   The nucleic acid sequence according to
claim 64 wherein said nucleic acid comprises the
nucleotide sequence set forth in Fig. 38C.

66.   An isolated protein comprising a CF or CFI
form of the amino acid sequence set forth in any one
of Figs. 39A-127A.

67.   A nucleic acid comprising the nucleotide
sequence set forth in Fig. 39B.

68.   A nucleic acid comprising the nucleotide
sequence set forth in Fig. 40B.

69.   A nucleic acid comprising the nucleotide
sequence set forth in Fig. 41B.

70.   A nucleic acid comprising the nucleotide
sequence set forth in Fig. 42B.

71.   A nucleic acid comprising the nucleotide
sequence set forth in Fig. 43B.

72.   A nucleic acid comprising the nucleotide
sequence set forth in Fig. 44B.

73. A nucleic acid comprising the nucleotide sequence set forth in Fig. 45B.

74. A nucleic acid comprising the nucleotide sequence set forth in Fig. 46B.

75. A nucleic acid comprising the nucleotide sequence set forth in Fig. 47B.

76. A nucleic acid comprising the nucleotide sequence set forth in Fig. 48B.

77. A nucleic acid comprising the nucleotide sequence set forth in Fig. 49B.

78. A nucleic acid comprising the nucleotide sequence set forth in Fig. 50B.

79. A nucleic acid comprising the nucleotide sequence set forth in Fig. 51B.

80. A nucleic acid comprising the nucleotide sequence set forth in Fig. 52B.

81. A nucleic acid comprising the nucleotide sequence set forth in Fig. 53B.

82. A nucleic acid comprising the nucleotide sequence set forth in Fig. 54B.

83. A nucleic acid comprising the nucleotide sequence set forth in Fig. 55B.

84. A nucleic acid comprising the nucleotide sequence set forth in Fig. 56B.

85.   A nucleic acid comprising the nucleotide
sequence set forth in Fig. 57B.

86.   A nucleic acid comprising the nucleotide
sequence set forth in Fig. 58B.

87.   A nucleic acid comprising the nucleotide
sequence set forth in Fig. 59B.

88.   A nucleic acid comprising the nucleotide
sequence set forth in Fig. 60B.

89.   A nucleic acid comprising the nucleotide
sequence set forth in Fig. 61B.

90.   A nucleic acid comprising the nucleotide
sequence set forth in Fig. 62B.

91.   A nucleic acid comprising the nucleotide
sequence set forth in any one of Figs. 63B-84B, 65D,
67D and 68D.

92.   A nucleic acid comprising the nucleotide
sequence set forth in any one of Figs. 85B-106B,
88D, 90D and 92D.

93.   A nucleic acid comprising the nucleotide
sequence set forth in any one of Figs. 107B-127B,
109D, 111D and 112D.

94.   A vector comprising the nucleic acid
according to any one of claims 2-7, 9-28, 31, 32,
34, 35, 37, 38, 40, 41, 43, 44, 46, 47, 49, 50, 52,
53, 55, 56, 58, 59, 61, 62, 64, 65 and 67-93.

95. A composition comprising at least one protein or nucleic acid according to any one of claims 1-93 and a carrier.

96. A method of inducing an immune response in a mammal comprising administering to said mammal an amount of at least one protein and/or nucleic acid according to any one of claims 1-93 sufficient to effect said induction.

MRVMGIQRNCQHLWRWGTMILGMLMICSAAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDTEVHNVWAT

HACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTNVRNVSSNG

TETDNEEIKNCSFNITTELRDKKQKVYALFYRLDVVPIDDKNSSEISGKNSSEYYRLINCNTSAITQACP

KVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSEN

ITNNAKTIIVQLNESVEINCTRPNNNTRKSIHIGPGQAFYATGEIIGDIRQAHCNISRTKWNKTLQQVAK

KLREHFNNKTIIFKPSSGGDLEITTHSFNCGGEFFYCNTSGLFNSTWMFNGTYMFNGTKDNSETITLPCR

IKQIINMWQGVGQAMYAPPIEGKITCKSNITGLLLTRDGGNNSNKNKTETFRPGGGDMRDNWRSELYKYK

VVKIEPLGVAPTKAKRRVVEREKRAVGIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQSNLLR

AIEAQQHLLQLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTNVPWNSSWSNKSQDEIWDNMT

WMEWEREISNYTDIIYRLIEESQNQQEKNEQELLALDKWASLWNWFDITNWLWYIKIFIMIVGGLIGLRI

VFAVLSIVNRVRQGYSPLSFQTLIPNPRGPDRPEGIEEEGGEQGRDRSIRLVNGFLALAWDDLRSLCLFS

YHRLRDFILIAARTVELLGRRSLRGLQKGWEALKYLGNLLQYWGQELKNSAISLLDTTAIAVAEGTDRVI

EIVQRACRAILNIPRRIRQGLERALL

Fig. 1A

Fig. 1B

Cleavage site   Fusion domain   TM

gp160

gp140CF

gp120

Fig. 1C

gp120 gp140CF   MW

— 198 kd
— 115 kd
— 93 kd

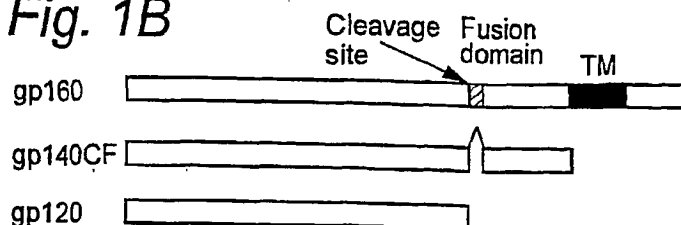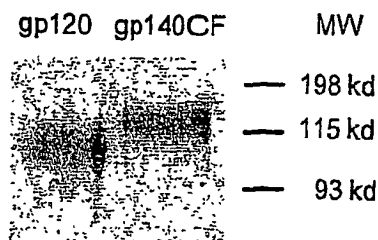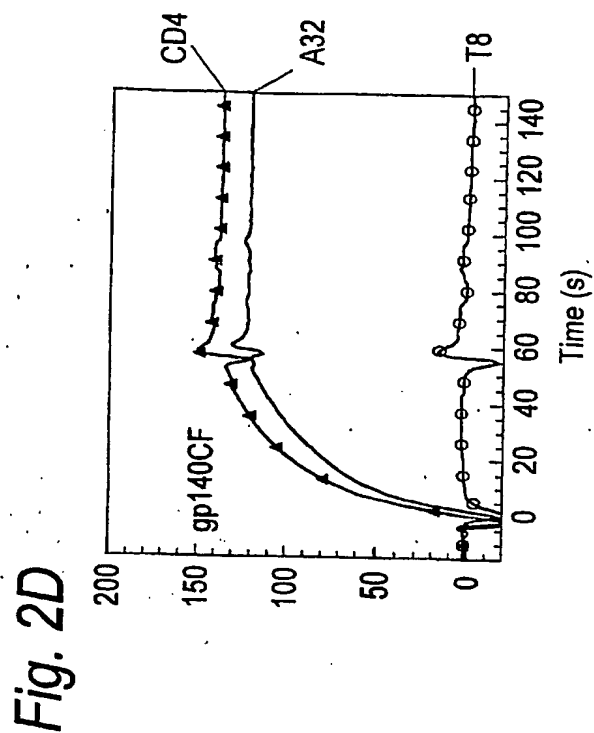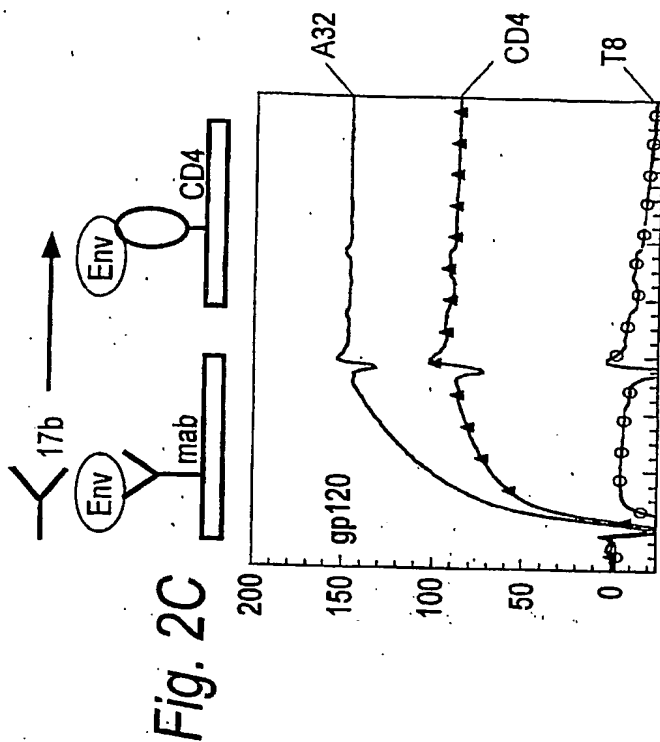## Fig. 1D

CON6.env (group M env consensus. This one contain five variable regions in env gene from 98CN006 virus, not in the public domain yet)

```
GCCACCATGCGCGTGATGGGCATCCAGCGCAACTGCCAGCACCTGTGGCGCTGGGGCACCATGATC
CTGGGCATGCTGATGATCTGCTCCGCCGCCGAGAACCTGTGGGTGACCGTGTACTACGGC
GTGCCCGTGTGGAAGGAGGCCAACACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTAC
GACACCGAGGTGCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCC
CAGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTG
GAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTCCCTGAAGCCCTGCGTGAAG
CTGACCCCCCTGTGCGTGACCCTGAACTGCACCAACGTGCGCAACGTGTCCTCCAACGGC
ACCGAGACCGACAACGAGGAGATCAAGAACTGCTCCTTCAACATCACCACCGAGCTGCGC
GACAAGAAGCAGAAGGTGTACGCCCTGTTCTACCGCCTGGACGTGGTGCCCATCGACGAC
AAGAACTCCTCCGAGATCTCCGGCAAGAACTCCTCCGAGTACTACCGCCTGATCAACTGC
AACACCTCCGCCATCACCCAGGCCTGCCCCAAGGTGTCCTTCGAGCCCATCCCCATCCAC
TACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACC
GGCCCCTGCAAGAACGTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGTCC
ACCCAGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGGAGATCATCATCCGCTCCGAGAAC
ATCACCAACAACGCCAAGACCATCATCGTGCAGCTGAACGAGTCCGTGGAGATCAACTGC
ACCCGCCCCAACAACAACACCCGCAAGTCCATCCACATCGGCCCCGGCCAGGCCTTCTAC
GCCACCGGCGAGATCATCGGCGACATCCGCCAGGCCCACTGCAACATCTCCCGCACCAAG
TGGAACAAGACCCTGCAGCAGGTGGCCAAGAAGCTGCGCGAGCACTTCAACAACAAGACC
ATCATCTTCAAGCCCTCCTCCGGCGGCGACCTGGAGATCACCACCCACTCCTTCAACTGC
GGCGGCGAGTTCTTCTACTGCAACACCTCCGGCCTGTTCAACTCCACCTGGATGTTCAAC
GGCACCTACATGTTCAACGGCACCAAGGACAACTCCGAGACCATCACCCTGCCCTGCCGC
ATCAAGCAGATCATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTACGCCCCCCCCATC
GAGGGCAAGATCACCTGCAAGTCCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGC
AACAACTCCAACAAGAACAAGACCGAGACCTTCCGCCCCGGCGGCGGCGACATGCGCGAC
AACTGGCGCTCCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCC
CCCACCAAGGCCAAGCGCCGCGTGGTGGAGCGCGAGAAGCGCGCCGTGGGCATCGGCGCC
GTGTTCCTGGGCTTCCTGGGCGCCGCCGGCTCCACCATGGGCGCCGCCTCCATCACCCTG
ACCGTGCAGGCCCGCCAGCTGCTGTCCGGCATCGTGCAGCAGCAGTCCAACCTGCTGCGC
GCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAG
GCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGC
TGCTCCGGCAAGCTGATCTGCACCACCAACGTGCCCTGGAACTCCTCCTGGTCCAACAAG
TCCCAGGACGAGATC TGGGACAACATGACCTGGATGGAGTGGGAGCGCGAGATCTCCAAC
TACACCGACATCATCTACCGCCTGATCGAGGAGTCCCAGAACCAGCAGGAGAAGAACGAG
CAGGAGCTGCTGGCCCTGGACAAGTGGGCCTCCCTGTGGAACTGGTTCGACATCACCAAC
TGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGATCGGCCTGCGCATC
GTGTTCGCCGTGCTGTCCATCGTGAACCGCGTGCGCCAGGGCTACTCCCCCCTGTCCTTC
CAGACCCTGATCCCCAACCCCCGCGGCCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGC
GGCGAGCAGGGCCGCGACCGCTCCATCCGCCTGGTGAACGGCTTCCTGGCCCTGGCCTGG
GACGACCTGCGCTCCCTGTGCCTGTTCTCCTACCACCGCCTGCGCGACTTCATCCTGATC
GCCGCCCGCACCGTGGAGCTGCTGGGCCGCCGCTCCCTGCGCGGCCTGCAGAAGGGCTGG
GAGGCCCTGAAGTACCTGGGCAACCTGCTGCAGTACTGGGGCCAGGAGCTGAAGAACTCC
GCCATCTCCCTGCTGGACACCACCGCCATCGCCGTGGCCGAGGGCACCGACCGCGTGATC
GAGATCGTGCAGCGCGCCTGCCGCGCCATCCTGAACATCCCCCGCCGCATCCGCCAGGGC
CTGGAGCGCGCCCTGCTGTAA
```
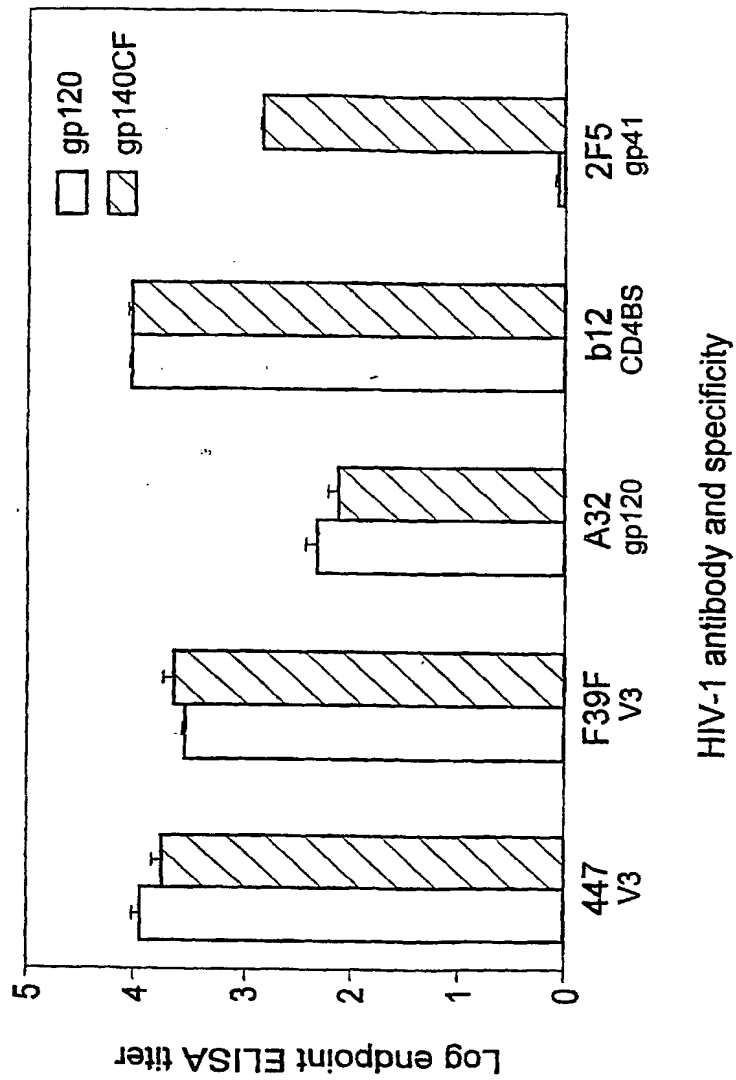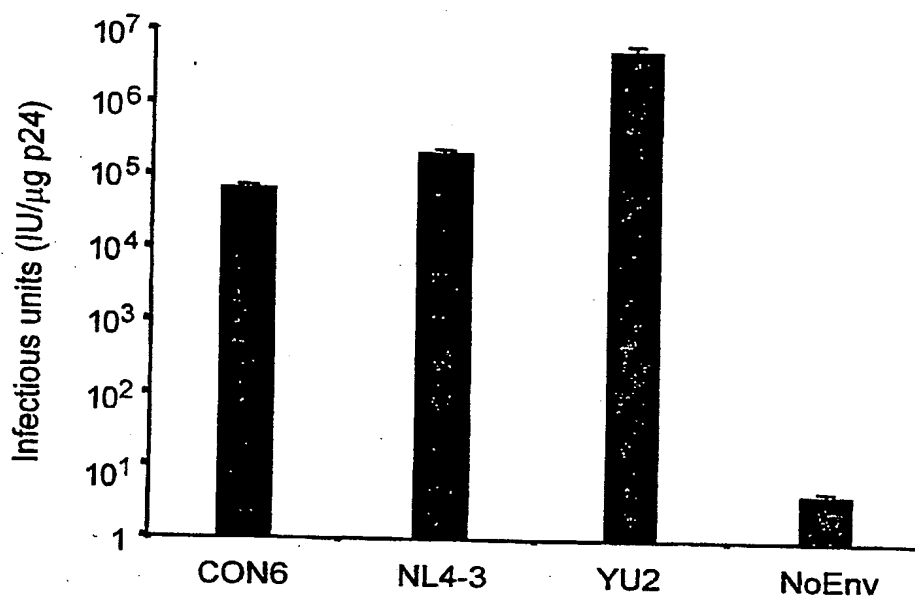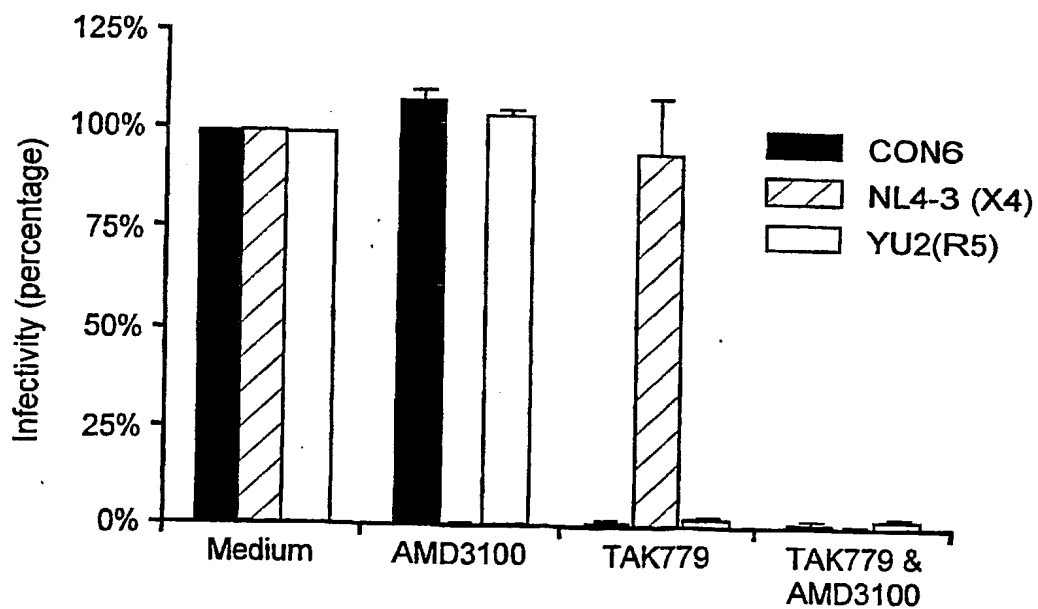
Fig. 2A

Fig. 2B

Fig. 2C

Fig. 2D

Fig. 2E

Fig. 3A

Fig. 3B

Fig. 4

Fig. 5

Fig. 6A

C.anc.env (subtype C ancestral env. The amino acid sequence is different from Los Alamos Database August 2002)

GCCG CCATG CG CGTGA TGGGCAT CC TGCGCAACTG CCAGCAG TGG TG GAT
CTGG GG CAT CC TGGGC TTC TGGA TG CTGATGAT CTGCTCC GT GGT GG GCA
ACC TGT GGG TGA CCG TG TACTAC GG CGT GC CCG TG TGGAA GGAGG CC AAG
ACC ACC CTG TT CTG CG CCTCC GA CG CCAAGG CCTA CGAGC GG CGAGG T GCA
CAAC GTG TG GG CCA CC CA CGC CTG CGTG CC CACC GA CCC CAA CCC CC AGG
AGAT GG TGC TG GAGAA CG TGA CC GA GAA CT TCAACAT GTG GAAGAA AC GAC
ATG GT GGA CCA GAT GCACG AGG A CA TCA TC TCCC TG TGGGA CCAG TC CCT
GAAG CC CTG CG TGAAG CTGACC CC CC TG TG CGT GA CC CTGAA CTG CA CCA
ACGT GA CCAAC GCCAC CAAC AAC ACC CTA CAACG GC GAGAT GAAGAAC TGC
TCC TT CAACAT CAC CA CCGAG CT GC GCGA CAAGAA GA AGAAG GAG TA CG C
CCT GTT CTA CC GCC TGGA CAT CG TG CCC CT GAA CGAGAAC TC CTC CG AGT
ACCG CC TGATCAACTG CAACA CC TC CGC CA TCACC CAGGC CTG CC CC AAG
GTGT CC TTC GA CCCC CA TCCCC CAT CCA CTAC TGC GC CCC CCG CCGGC TA CGC
CATC CT GAAG GT GCAACAA CAA GA CC TTC AAC GG CA CCGGC CCC TG CA ACA
ACGT GT CCA CC GTG CA GTG CA CC CA CGG CAT CAAG CC CGT GG TGT CC ACC
CAG CTG CTG CT GAAC GG CTCC CT GGC CGAGG AGGA GA TCA TCA TCC CG CT C
CGA GAAC CT GA CC GA CAACGC CAAGACC AT CAT CG TG CAG CT GAA CG AGT
CCGT GGAGA TC GTGTG CA CCC GC CCC CAA CA ACA ACA CC CG CAAGT CC ATG
CG CATC GGC CCC GGC CCA GAC CTT CTA CG CCA CCGG CGA CA TCA TCGG CGA
CATC CG CCA GG CCCA CTGC AA CA TC TCC GA GGA CAAG TGGAA CAA GA CC C
TG CAGCAGG TG GCC GA GAAG CTG GGG CAA GCA CTTC CCC AA CAAGA CC AT C
ACC TTC GAGC CCT CC T CCGGC GG CGA CC TG GAGAT CA CCA CC CACTC CTT
CAAC TGC CCG CG GCGAG TTC TTC TAC TGC AA CACC T CCAAGC TG TT CAA CT
CCAC CTAC AA CAAC AA CA CC AAC TC CAAC T CCA CC AT CAC CC TGCC C TG C
CG CATCAAG CA GAT CA TCAACAT GTG GCAG GGC GT GGG CCAGG CC AT GTA
CG CC CC CCC CA TCG CC GG CCAA CA TCACC TGC AAG T CCAA CAT CAC CG GC C
TG CT GCTGA CC CGC GA CG GC GG CAAG GA GAA CA CC AC CGA GA CC TTC CG C
CCC GG CGG CGG CGA CA TG CG CGA CAA CTG GCG CTC CGAG CTG TACAA GTA
CAAG GT GGT GG AGA TCAAGCC CC TG GGC GT GGC CCC CCACC GA GGC CA AG C
GCCG CGTGG TG GAG CG CGA GAAG CG CGC CG TG GGC CTG GG CG CCG TG TT C
CTGGG CTTC CT GGGCG CC GCC GG CT CCA CCATG GG CGC CG CC TCC AT CAC
CCT GA CCG TG CA GG CC CG CCA GC TG CTG TC CGG CA TCG TG CAG CAG CAG T
CCAA CCT GC TG CGC GC CAT CGA GG CC CAG CAG CACAT GCT GCAG CT GA CC`
GT GT GGG GCAT CAAG CAG CTG CA GG CCC GC GT GCT GG CCA TG GAG CG CTA
CCT GAAG GA CCAG CAG CTG CT GG GCATC TG GGG CT GCT CC GG CAAG C TGA
TCTG CA CCA CC GCC GTG CCC CT GGA A CTC CT CCT GG TC CAA CA AGT CC CTG
GAC GACAT CTG GGA CAA CAT GAC CT GGA TG GAG TG GGA CC GC GAGAT CT C
CAAC TA CAC CGA CA CCA TCTA CC GC CTG CT GGA GGA GTC C CA GAA CC AG C
AGGA GAAGAA C GAG CAG GAC CTG CT GG CC CT GGA CT CC TG GGA GAA C TG
TGGAA CTGG TT CGA CA TCACC AA CT GG CTG TGG TA CAT CA AGA TC TT CAT
CAT GAT CG TG GG CG GC CTGAT CG GC CTG CG CAT CA TC TTC GC CG TG C TG T
CCAT CG TGAA CC GC GT GC GC CAG GG CTA CT CCC CC CTG TC CT TCCAG AC C
CTGA CC CCC AA CCC CC GCGGC CC CGA CC GC CTG GA GCGCA TC GAG GA GGA
GGGC GG CGA GCAG GAC CGC GA CC GC TCC AT CCG CC TG GTG TC CGG CT TC C
TGGC CCT GG CC TGG GA CGA CC TG CG CTC CCT GTG CC TG TT CT CCTAC CAC
CG CC TG CGC GA CTT CA TCCT GAT CG CCG CC CGC AC CG TG GAG CTG CTG GG
CCG CTC CT CCC TG CG CG GC CT GCAG CG CGG CTG GGA GG CC CT GAA GTA CC
TGGG CT CCC TG GTG CA GTA CT GG GG CCA GGA GC TGA AGAA GT CCG CC AT C
TCCCT GC TGGA CA CCA TCG CC AT CG CCG TG GCC GA GGG CA CC GA CCG CAT
CAT CGAGG TG GTG CAG CG CGC CT GC CG CGC CAT CC TGAA CAT CCC CCC GCC
GCATCCG C~~C~~TA~~ERC~~ ~~CT~~ T~~tcT~~ A~~CEc~~ .~~CeHEETrc~~f~~R~~dTiA~~X~~

Fig. 6B

C.con.env (subtype C consensus env. The amino acid sequence
is different from Los Alamos Database August 2002)

GCCGCCATGCGCGTGATGGGCATCCTGCGCAACTGCCAGCAGTGGTGGAT
CTGGGGCATCCTGGGCTTCTGGATGCTGATGATCTGCAACGTGGTGGGCA
ACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAAG
ACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGAGAAGG AGGTGCA
CAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGG
AGATGGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACGAC
ATGGTGGACCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTCCCT
GAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGAACTGCCGCA
ACGTGACCAACGCCACCAACAACACCTACAACGAGGAGATCAAG AACTGC
TCCTTCAACATCACCACCGAGCTGCGCGACAAGAAGAAGAAGGTGTACGC
CCTGTTCTACCGCCTGGACATCGTGCCCCTGAACGAGAACTCCTCCGAGT
ACCGCCTGATCAACTGCAACACCTCCGCCATCACCCAGGCCTGCCCCAAG
GTGTCCTTCGACCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTACGC
CATCCTGAAGTGCAACAACAAGACCTTCAACGGCACCGGCCCCTG CAACA
ACGTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGTCCACC
CAGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGGAGATCATCATCCGCTC
CGAGAACCTGACCAACAACGCCAAGACCATCATCGTGCACCTGAACGAGT
CCGTGGAGATCGTGTGCACCCGCCCCAACAACAACACCCGCAAGTCCATC
CGCATCGGCCCCGGCCAGACCTTCTACGCCACCGGCGACATCATCG GCGA
CATCCGCCAGGCCCACTGCAACATCTCCGAGGACAAGTGGAACAAGACCC
TGCAGCGCGTGTCCAAGAAGCTGAAGGAGCACTTCCCCAACAAGACCATC
AAGTTCGAGCCCTCCTCCGGCGGCGACCTGGAGATCACCACCCACTCCTT
CAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCAAGCTGTTCAACT
CCACCTACAACAACAACACCAACTCCAACTCCACCATCACCCTGCCC TGC
CGCATCAAGCAGATCATCAACATGTGGCAGGAGGTGGGCCGCGCCATGTA
CGCCCCCCCCATCGCCGGCAACATCACCTGCAAGTCCAACATCACCGGCC
TGCTGCTGACCCGCGACGGCGGCAAGAAGAACACCACCGAGATCTTCCGC
CCCGGCGGCGGCGACATGCGCGACAACTGGCGCTCCGAGCTGTACAAGTA
CAAGGTGGTGGAGATCAAGCCCCTGGGCGTGGCCCCCACCAAGGCCAA GC
GCCGCGTGGTGGAGCGCGAGAAGCGCGCCGTGGGCATCGGCGCCGTGTTC
CTGGGCTTCCTGGGCGCCGCCGGCTCCACCATGGGCGCCGCCTCCATCAC
CCTGACCGTGCAGGCCCGCCAGCTGCTGTCCGGCATCGTGCAGCAGCAGT
CCAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACATGCTGCAGCTGACC
GTGTGGGGCATCAAGCAGCTGCAGACCCGCGTGCTGGCCATCGAGCGCTA
CCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGCAAGCTGA
TCTGCACCACCGCCGTGCCCTGGAACTCCTCCTGGTCCAACAAGTCCCAG
GAGGACATCTGGGACAACATGACCTGGATGCAGTGGGACCGCGAGATCTC
CAACTACACCGACACCATCTACCGCCTGCTGGAGGACTCCCAGAACCAGC
AGGAGAAGAACGAGAAGGACCTGCTGGCCCTGGACTCCTGGAAGAACCTG
TGGAACTGGTTCGACATCACCAACTGGCTGTGGTACATCAAGATCTTCAT
CATGATCGTGGGCGGCCTGATCGGCCTGCGCATCATCTTCGCCGTGCTGT
CCATCGTGAACCGCGTGCGCCAGGGCTACTCCCCCCTGTCCTTCCAGACC
CTGACCCCCAACCCCCGCGGCCCCGACCGCCTGGGCCGCATCGAGGAGGA
GGGCGGCGAGCAGGACCGCGACCGCTCCATCCGCCTGGTGTCCGGCTTCC
TGGCCCTGGCCTGGGACGACCTGCGCTCCCTGTGCCTGTTCTCCTACCAC
CGCCTGCGCGACTTCATCCTGGTGGCCGCCCGCGCCGTGGAGCTGCTGGG
CCGCTCCTCCCTGCGCGGCCTGCAGCGCGGCTGGGAGGCCCTGAAGTACC
TGGGCTCCCTGGTGCAGTACTGGGGCCTGGAGCTGAAGAAGTCCGCCATC
TCCCTGCTGGACACCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCAT
CATCGAGCTGATCCAGCGCATCTGCCGCGCCATCCGCAACATCCCCCGCC
GCATCCG~~CCAGGGGTTGGAGCGGCCTTCAGTAA~~

C.anc.env (subtype C ancestral env)

```
MRVMGILRNCQQWWIWGILGFWMLMICSVVGNLWVTVYYGVPVWKEAKTTLFCASDAKAYEREVHNVWAT
HACVPTDPNPQEMVLENVTENFNMWKNDMVDQMHEDIISLWDQSLKPCVKLTPLCVTLNCTNVTNATNNT
YNGEMKNCSFNITTELRDKKKKEYALFYRLDIVPLNENSSEYRLINCNTSAITQACPKVSFDPIPIHYCA
PAGYAILKCNNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNAKTIIVQLN
ESVEIVCTRPNNNTRKSMRIGPGQTFYATGDIIGDIRQAHCNISEDKWNKTLQQVAEKLGKHFPNKTITF
EPSSGGDLEITTHSFNCRGEFFYCNTSKLFNSTYNNTNSNTITLPCRIKQIINMWQGVGQAMYAPPIA
GNITCKSNITGLLLTRDGGKENTTETFRPGGGDMRDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKR
AVGLGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVL
AMERYLKDQQLLGIWGCSGKLICTTAVPWNSSWSNKSLDDIWDNMTWMEWDREISNYTDTIYRLLEESQN
QQEKNEQDLLALDSWENLWNWFDITNWLWYIKIFIMIVGGLIGLRIIFAVLSIVNRVRQGYSPLSFQTLT
PNPRGPDRLERIEEEGGEQDRDRSIRLVSGFLALAWDDLRSLCLFSYHRLRDFILIAARTVELLGRSSLR
GLQRGWEALKYLGSLVQYWGQELKKSAISLLDTIAIAVAEGTDRIIEVVQRACRAILNIPRRIRQGFEAA
LL
```

Fig. 6C

C.con.env (subtype C consensus env)

```
MRVMGILRNCQQWWIWGILGFWMLMICNVVGNLWVTVYYGVPVWKEAKTTLFCASDAKAYEKEVHNVWAT
HACVPTDPNPQEMVLENVTENFNMWKNDMVDQMHEDIISLWDQSLKPCVKLTPLCVTLNCRNVTNATNNT
YNEEIKNCSFNITTELRDKKKVYALFYRLDIVPLNENSSEYRLINCNTSAITQACPKVSFDPIPIHYCA
PAGYAILKCNNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTNNAKTIIVHLN
ESVEIVCTRPNNNTRKSIRIGPGQTFYATGDIIGDIRQAHCNISEDKWNKTLQRVSKKLKEHFPNKTIKF
EPSSGGDLEITTHSFNCRGEFFYCNTSKLFNSTYNNTNSNTITLPCRIKQIINMWQEVGRAMYAPPIA
GNITCKSNITGLLLTRDGGKNNTEIFRPGGGDMRDNWRSELYKYKVVEIKPLGVAPTKAKRRVVEREKR
AVGIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQTRVL
AIERYLKDQQLLGIWGCSGKLICTTAVPWNSSWSNKSQEDIWDNMTWMQWDREISNYTDTIYRLLEDSQN
QQEKNEKDLLALDSWKNLWNWFDITNWLWYIKIFIMIVGGLIGLRIIFAVLSIVNRVRQGYSPLSFQTLT
PNPRGPDRLGRIEEEGGEQDRDRSIRLVSGFLALAWDDLRSLCLFSYHRLRDFILVAARAVELLGRSSLR
GLQRGWEALKYLGSLVQYWGLELKKSAISLLDTIAIAVAEGTDRIIELIQRICRAIRNIPRRIRQGFEAA
LQ
```

Fig. 6D

Fig. 6E

Synthesize entire gene in 80-mer fragments overlapping by 20 residues
at the 3' end with invariant sequences at the 5' end.

EcoRI   BbsI

80mer fragment

BsmBI   BamHI

80mer fragment

Invariant Seq.
(18 nt)
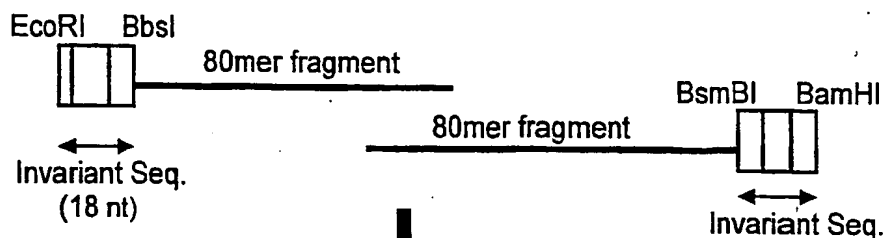
Invariant Seq.

Paired 80mer oligos are connected via PCR in a stepwise manner from 5' to 3' using
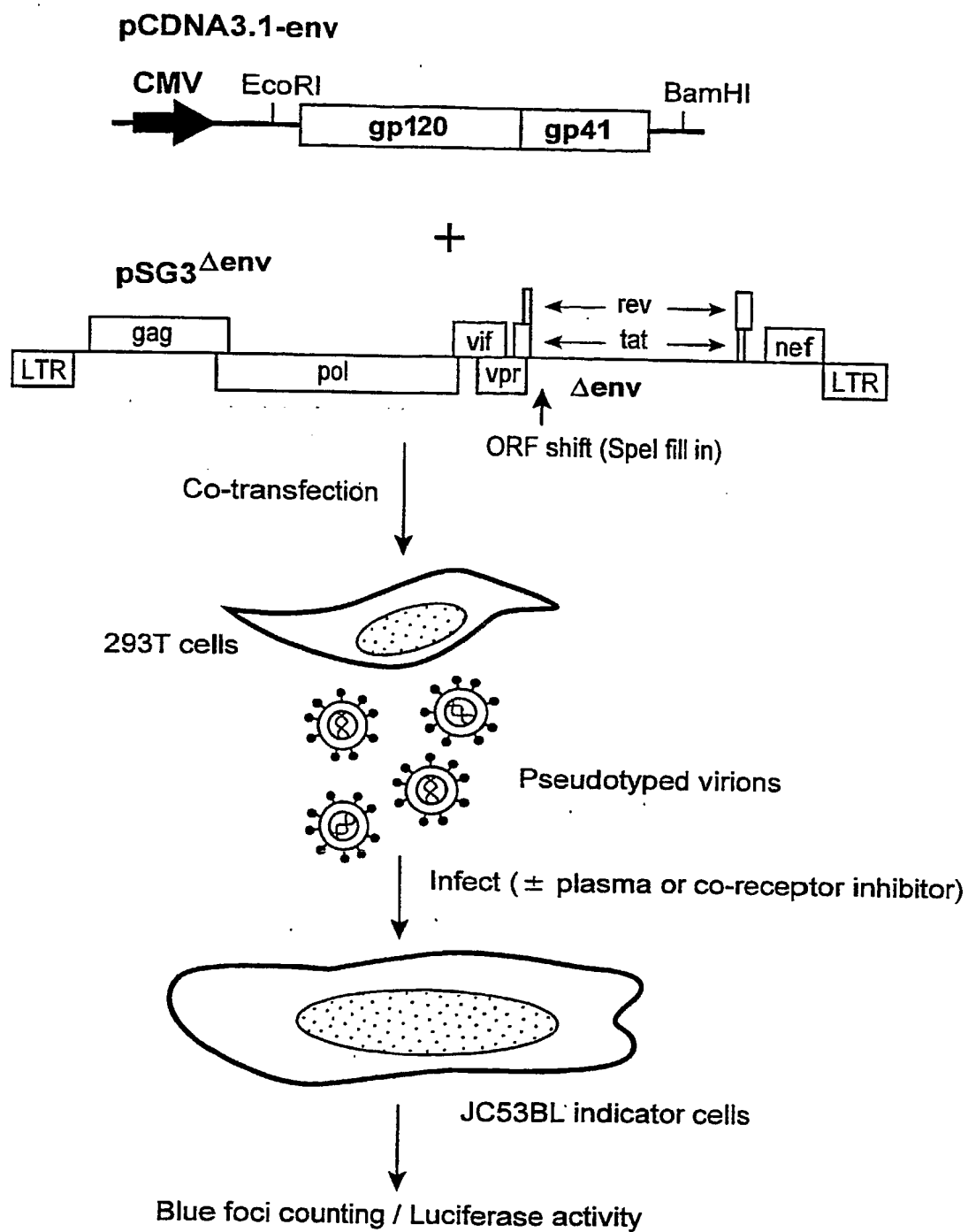primers complimentary to the invariant seq.

80mer Fragment

5'primer   140bp PCR product   3'primer
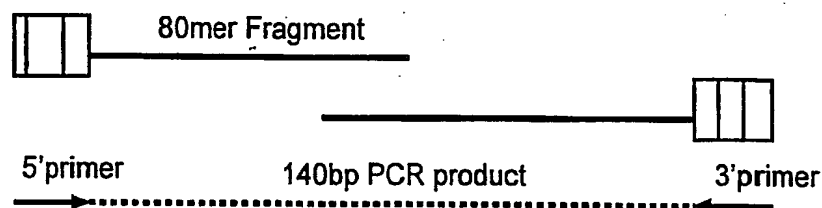
108bp PCR fragments cloned into pGEM-T and sequenced. Clones with the proper sequence
will be cut with 2 restriction enzymes. 4 fragments will be ligated together with pcDNA3.1 in
a stepwise manner from the 5' to 3' end of gene

Fragments to be ligated with pcDNA3.1 (1-4 are in order from 5' to 3')	Restriction Enzymes Used to Cleave Fragment
Fragment 1	EcoRI/BsmBI
Fragment 2	BbsI/BsmBI
Fragment 3	BbsI/BsmBI
Fragment 4	BbsI/BamHI
pcDNA3.1	EcoRI/BamHI

Fragment 2

Fragment 1   Fragment 3

EcoRI   Fragment 4

Gene
constructed
in pcDNA3.1   BamHI
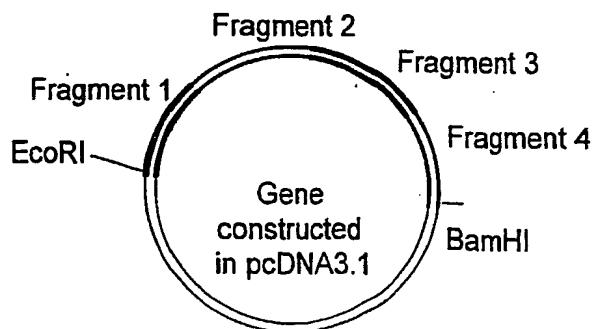
Ligations will be repeated stepwise 5' to 3' until the entire gene
has been cloned into pcDNA3.1
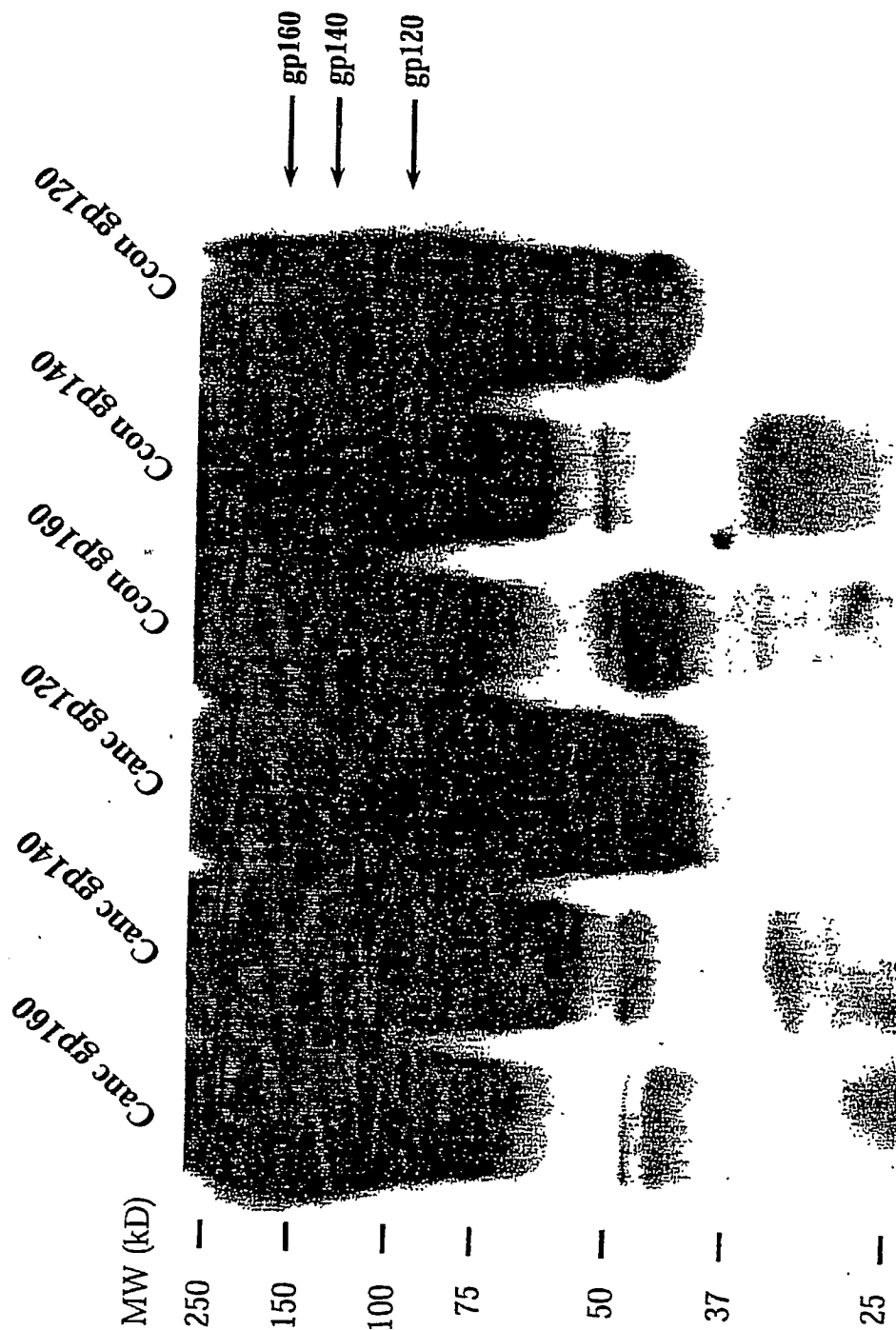
pCDNA3.1-env

CMV   EcoRI   gp120   gp41   BamHI

+

pSG3$^{\Delta env}$

LTR   gag   pol   vif   vpr   rev   tat   nef   LTR

ORF shift (SpeI fill in)

Δenv

Co-transfection

293T cells

Pseudotyped virions

Infect ($\pm$ plasma or co-receptor inhibitor)

JC53BL indicator cells

Blue foci counting / Luciferase activity

Fig. 7

Fig. 8

Fig. 9

Fig. 10A

Fig. 10B

Fig. 11

Fig. 12A

Fig. 12B

Fig. 12C

Fig. 13A

Gag

Fig. 13B

Nef
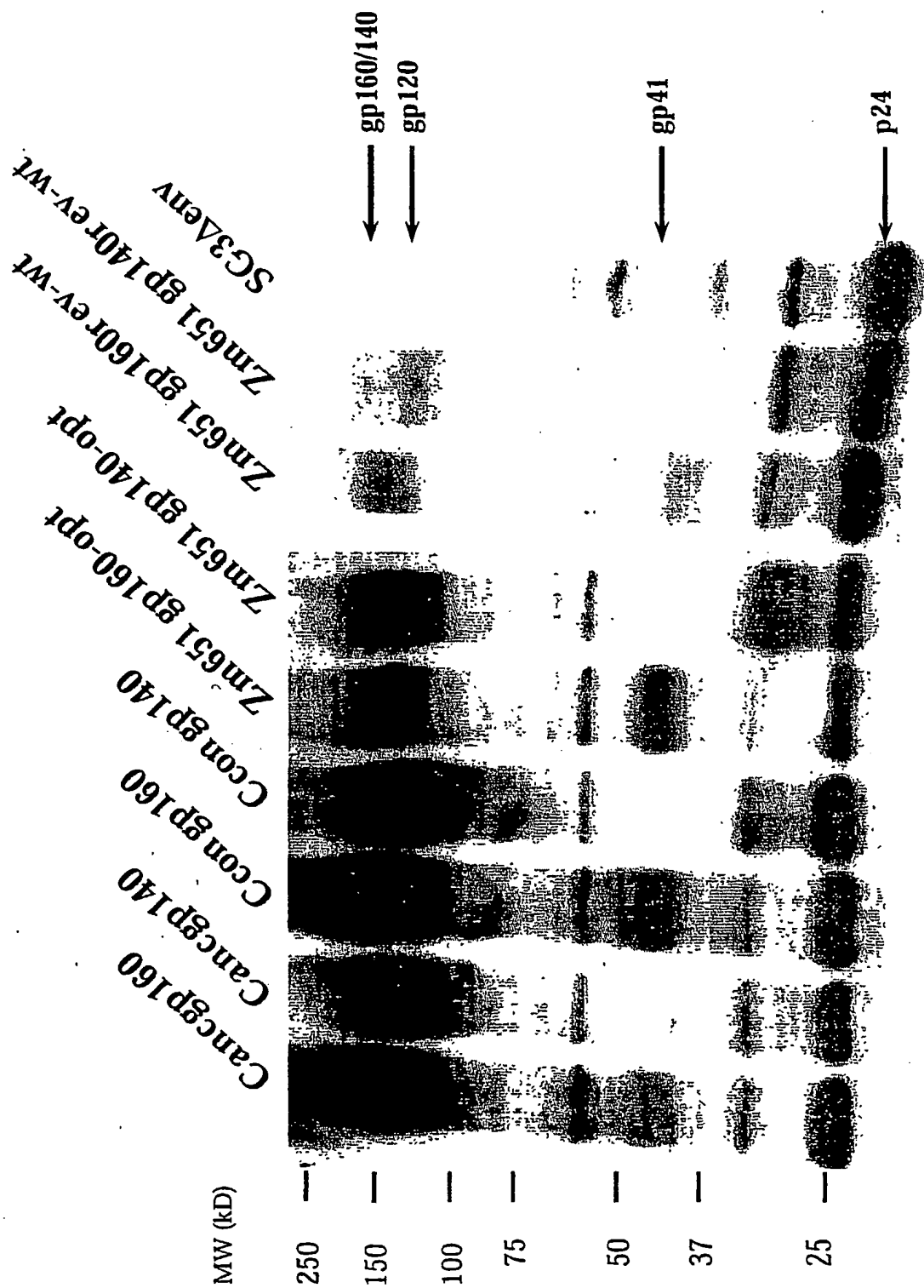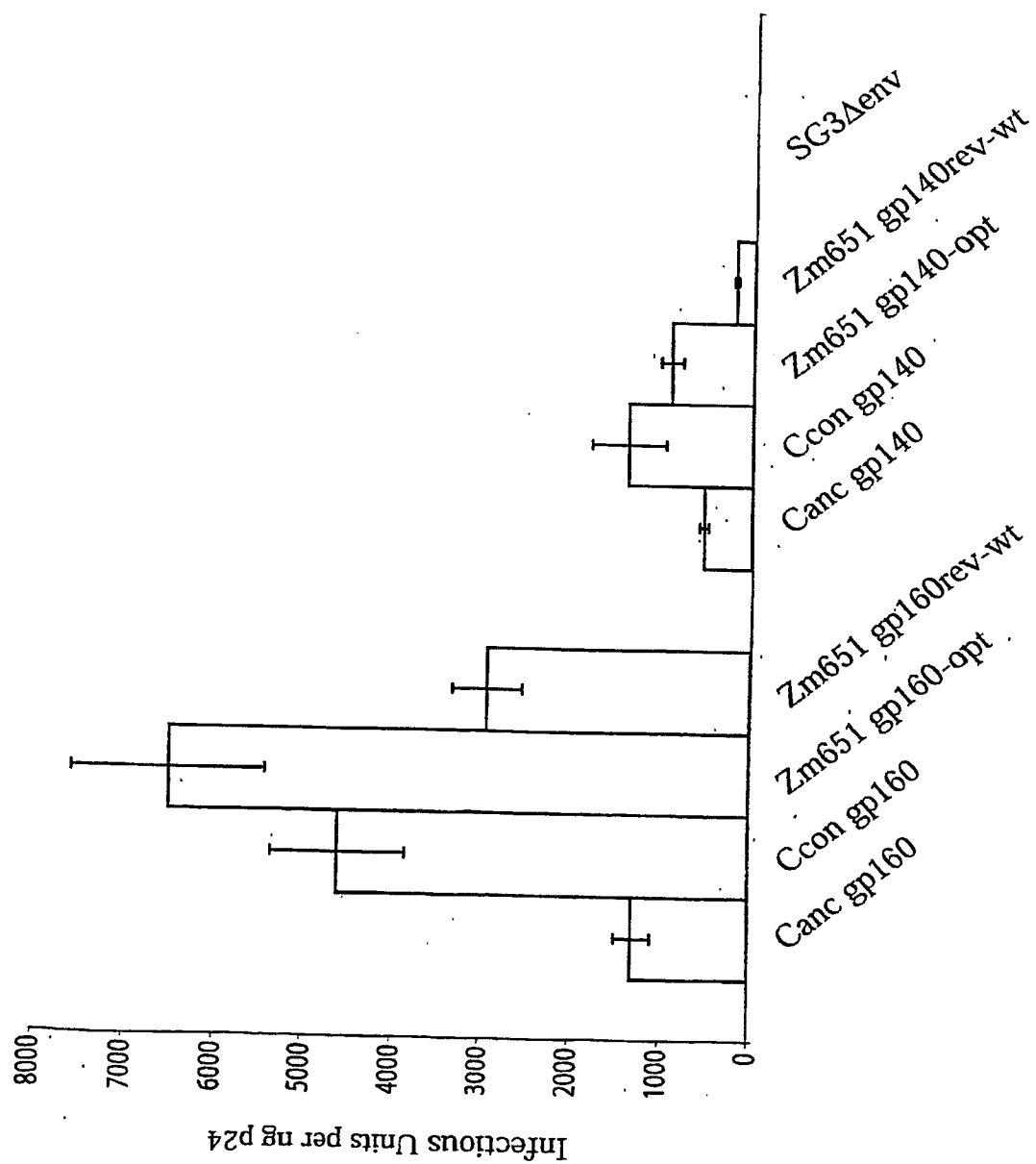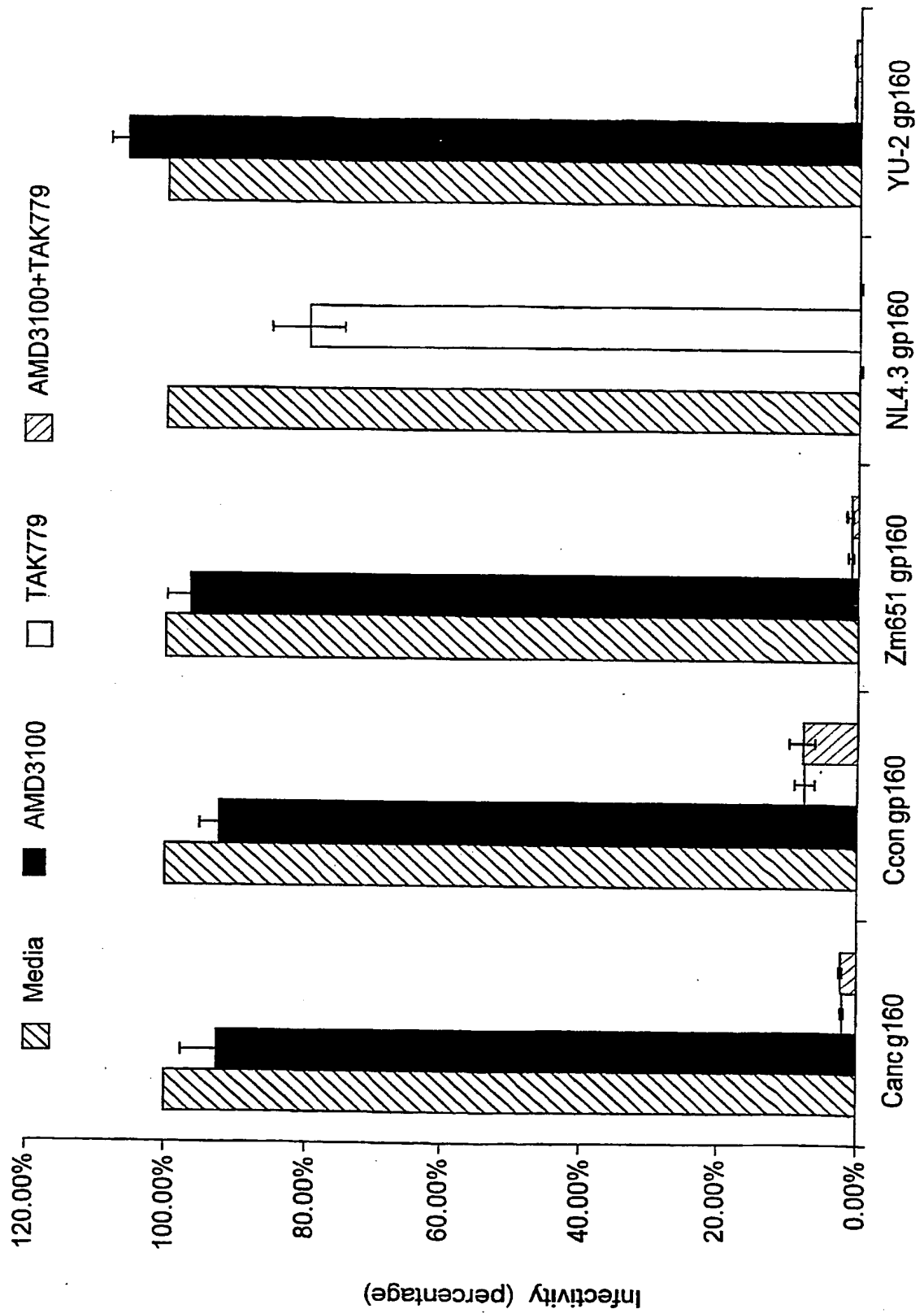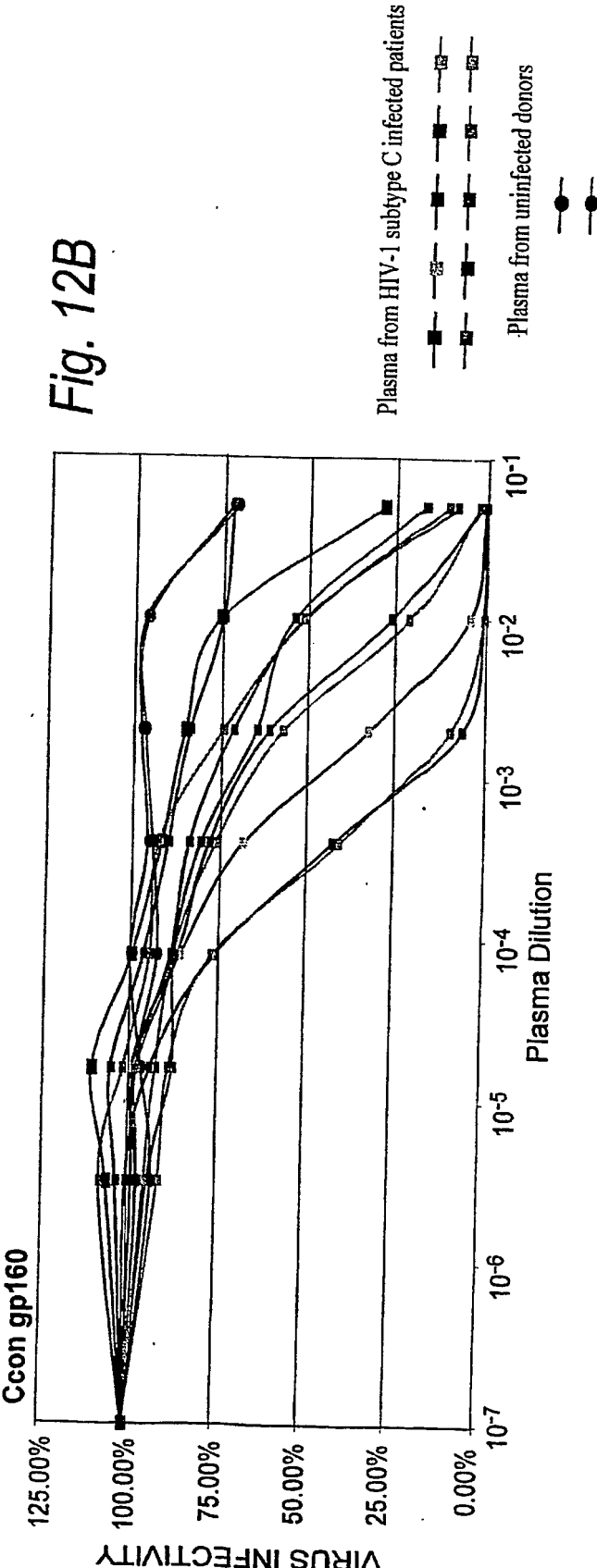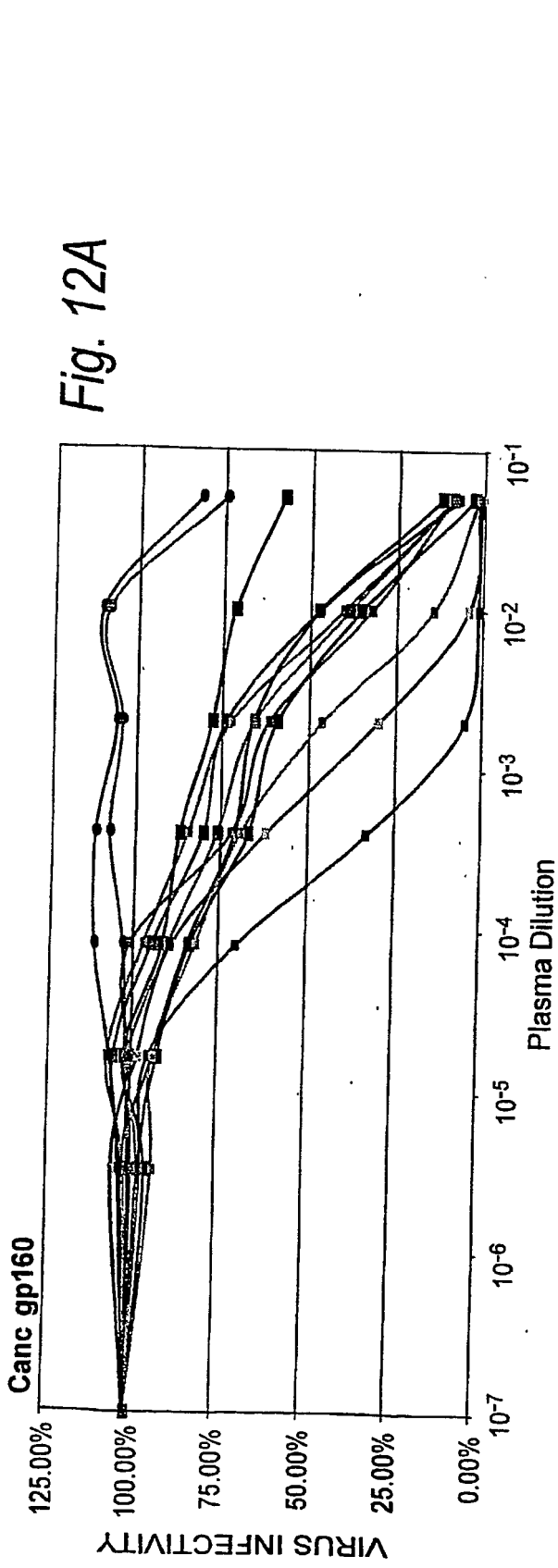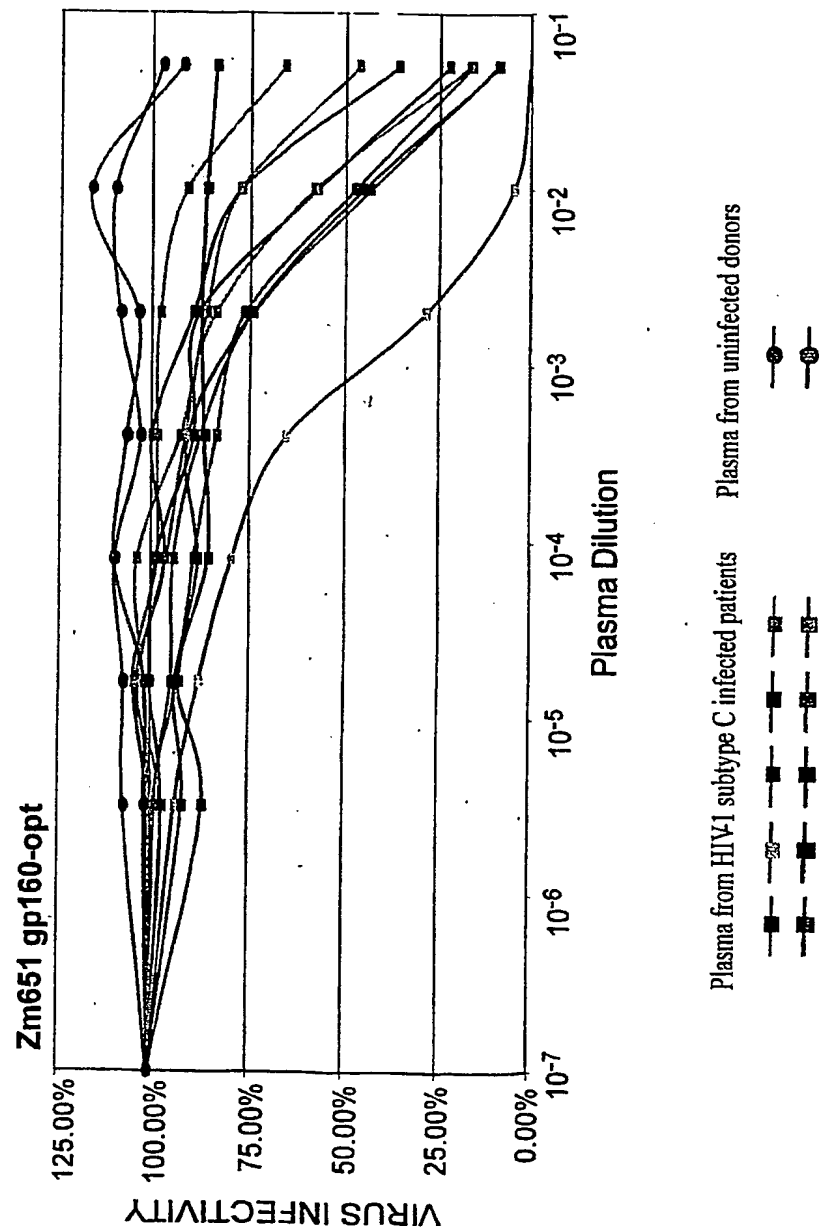
Fig. 13C
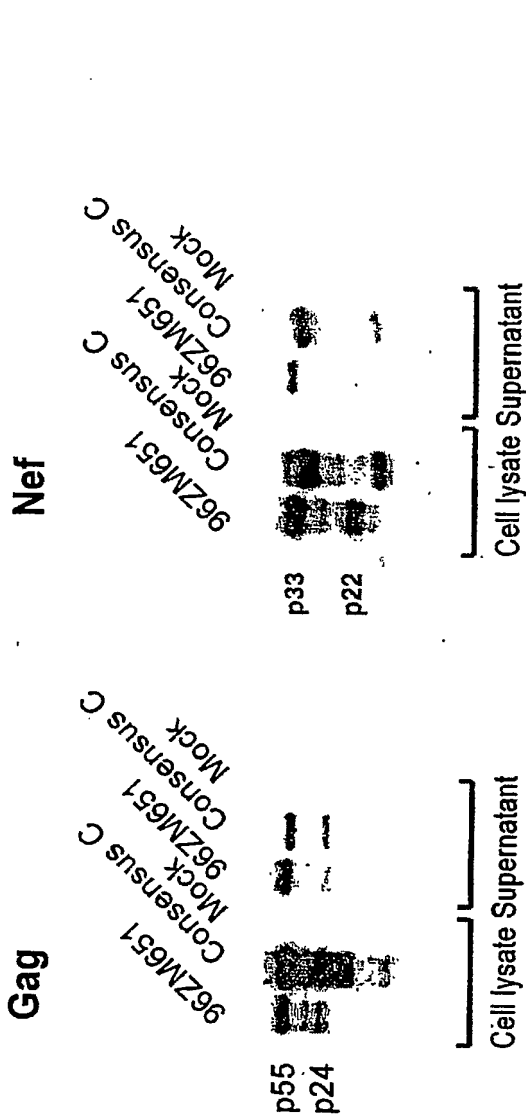
C.con.gag (subtype C con sensus gag)

MGARASILRGGKLDTWEKIRLRPGGKKRYMIKHLVWASRELERFALNPGLLETSEGCKQIMKQLQPA
LQTGTEELRSLYNTVATLYCVHEKIEVRDTKEALDKIEEEQNKSQQKTQQAEAAADGKVSQNYPI
VQNLQGQMVHQAISPRTLNAWVKVIEEKAFSPEVIPMFTALSEGATPQDLNTMLNTVGGHQAAMQMLKDT
INEEAAEWDRLHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIAWMTSNPPVPVGDIYKRWIILGLNKIV
RMYSPVSILDIKQGPKEPFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQNANPDCKTILRALGPGASLE
EMMTACQGVGGPSHKARVLAEAMSQANNTNIMMQRSNFKGPKRIVKCFNCGKEGHIARNCRAPRKKGCWK
CGKEGHQMKDCTERQANFLGKIWPSHKGRPGNFLQSRPEPTAPPAESFRFEETTPA
PKQEPKDREPLTSLKSLFGSDPLSQ

Fig. 13D

C.con.nef (subtype C consensus nef)

MGGKWSKSSIVGWPAVRERIRRTEPAAEGVGAASQDLDKYGALTSSNTATNNADCAWLEAQEEEEV
GFPVRPQVPLRPMTYKAAFDLSFFLKEKGGLEGLIYSKKRQEILDLWVYHTQGFFPDWQNYTPGPGVRYP
LTFGWCFKLVPVDPREVEEANEGENNCLLHPMSQHGMEDEDREVLKWKFDSHLARRHMARELHPEYYKDC

C.con.gag (subtype C consensus gag. Not in the public domain)

GCCGCCGCCGCCATGGGCGCCCGGCCCGCCCGGCCCAGCATCCTCTGCGCGGCGGCGGCCAAGCTGGACACCTGGGAGAAGATCCGCC
TGCCCCCGGGCGGGCAAGAAGCGCTACATGATCAAGCACCTGGTGTGGGCCAGCCGCGAGCTGGAGCGCTT
CGCCCTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCCAGCAGATCATGAAGCAGCTGCAGCCCGCC
CTGCAGACCGGCACCGAGGAGCTGCGCAGCCTGTACAACACCGTGGCCACCCTGTACTGCGTGCACGAGA
AGATCGAGGTGCGCGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGAACAAGAGCCAGCAGAA
GACCCAGCAGGCCGCGAGGCCCGCGCCCAGCCCGGCCAGGTGCAGGTGAGCCACCCCATCGTGCAGAACCTGCAG
GGCCAGATGGTGCACCAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAGG
CCTTCAGCCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAGGACCTGAACAC
CATGCTGAACACCGTGGGCGGCCACCAGGCCATCAGCCGCATGCTGAAGGACACCATCAACGAGGAGGCC
GCCGAGTGGGACCGCCTGCACCCCGTGCACGCGGGCCCCATCGCCCCGGCCAGATGCGCGAGCCCCGCG
GCAGCGACATCGCCGGCACCACCAGCACCCTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCGT
GCCCGTGGGCGACATCTACAAGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACAGCCCC
GTGAGCATCCTGGACATCAAGCAGGGCCCCAAGGAGCCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGA
CCCTGCGCGCCGAGCAGGCCACCCAGGACGTGAAGAACTGGATGACCGACACCCTGCTGGTGCAGAACGC
CAACCCCGACTGCAAGACCATCCTGCGCGCCCTGGGCCCGGTGCTGCCGAGGCCATGAGCCAGGCCAACACA
TGCCAGGGCGTGGGCGGCTGCCCCAGCCACAAGGCCGCGCAACTGCCGCGGCCACACTGCCCGGCGCAA
CCAACATCATGATGCAGCGCAGCAACTTCAAGGGCCCCATCCGCTGCTTCAACTGCGGCAA
GGAGGGCCACATCGCCCGCAACTGCCGCGAGCGCCAGGCCAAGAGGGGCTGCTGGAAGTGCGGCAAGGAGGGC
CACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCC
GCCCCGGCAACTTCCTGCAGAAGCGCCCCGAGCCCCACCGCGAGGCCTTCGCTTCGAGGA
GACCACCCCGCGCCCCCCAAGCAGGAGCCCAAGGAGCCCAGCCTGAAGAGCCTGTTCGGC
AGCGACCCCCTGAGCCAGTAA

Fig. 13E

C.con.nef (subtype C consensus nef. Not in the public domain)

GCCGCCGCCATGGGCGGCAAGTGGAGCAAGAGCAGCATCGTGGGCTGGCCCGCGGTGCGCGAGCGCATCC
GCCGCCACCGAGCCCGCGCCGCCGTGGGCCGCGCCGCCACCGACTGGACCTGGACAAGTACGGCGCCCTGAC
CAGCAGCAACACCGCCACCAACAACGCCGACTGCGCCCTGGCTGGAGGCCCAGGAGGAGGAGGTG
GGCTTCCCCGTGCGCCCCCAGGTGCCCCTGAGGCCCTACAAGGCCGCCTTCGACCTGAGCTTCT
TCCTGAAGGAGAAGGGCGGCCTGGAGGGCCTGATCTACAGCAAGAAGCGCCAGGAGATCCTGGACCTGTG
GGTGTACCACACCCAGGGCTTCTTCCCCGACTGGCAGAACTACACCCCCGGCCCCGGCGTGCGCTACCCC
CTGACCTTCGGCTGTGCTTCAAGCTGGTGCCCGTGGACCCCGAGGCCGTGGAGGAGGCCAACGAGGGCG
AGAACAACTGCCTGCTGCACCCCATGAGCCAGCACGGCATGGAGGACGAGGACCGCGAGGTGCTGAAGTG
GAAGTTCGACAGCAGCCTGGCGCCCGCCCGCACATGCCCGCGCCCCGAGTACTACAAGGACTGC
TGA

Fig. 13F

SUBSTITUTE SHEET (RULE 26)

CONs.env (gorup M consensus env gene. This one contain the consensus sequence for variable regions in env gene)

MRVRGIQRNCQHLWRWGTLILGMLMICSAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDTEVHNV
WATHACVPTDPNPQEIVLENVTENFNMWKNMVEQMHEDIISLMDQSLKPCVKLTPLCVTLNCTNVNVTN
TTNNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDVVPIDDNNNNSSNYRLINCNTSAITQACPKVSF
EPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIRSENITNN
AKTIIVQLNESVEINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNISGTKWNKTLQQVAKKLRE
HFNNKTIIFKPSSGGDLEITTHSFNCRGEFFYCNTSGLFNSTWIGNGTKNMNNTNDTITLPCRIKQIINM
WQGVGQAMYAPPIEGKITCKSNITGLLLTRDGGNNNTNETEIFRPGGGDMRDNWRSELYKYKVVKIEPLG
VAPTKAKRRVVEREKRAVGIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLRAIEAQQHL
LQLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTTVPWNSSWSNKSQDEIWDNMTWMEWEREI
NNYTDIIYSLIEESQNQQEKNEQELLALDKWASLWNWFDITNWLWYIKIFIMIVGGLIGLRIVFAVLSIV
NRVRQGYSPLSFQTLIPNPRGPDRPEGIEEEGGEQDRDRSIRLVNGFLALAWDDLRSLCLFSYHRLRDFI
LIAARTVELLGRKGLRRGWEALKYLWNLLQYWGQELKNSAISLLDTTAIAVAEGTDRVIEVVQRACRAIL
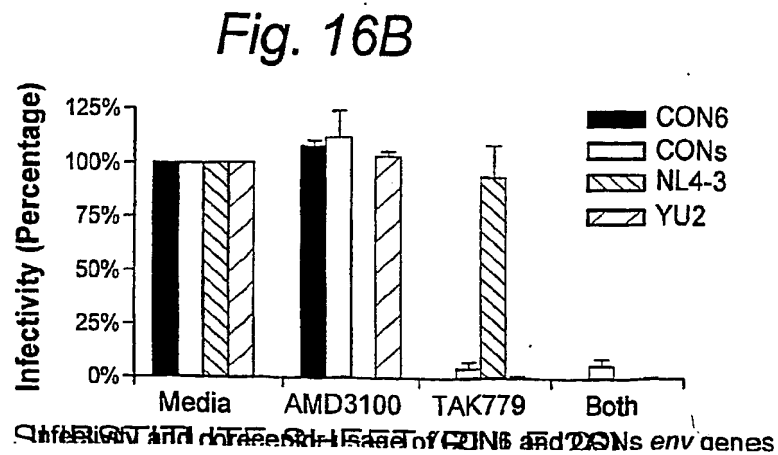NIPRRIRQGLERALL

Fig. 14A

Fig. 14C

gp160

CONS.gp160.1
CONS.gp160.2
CONS.gp160.3
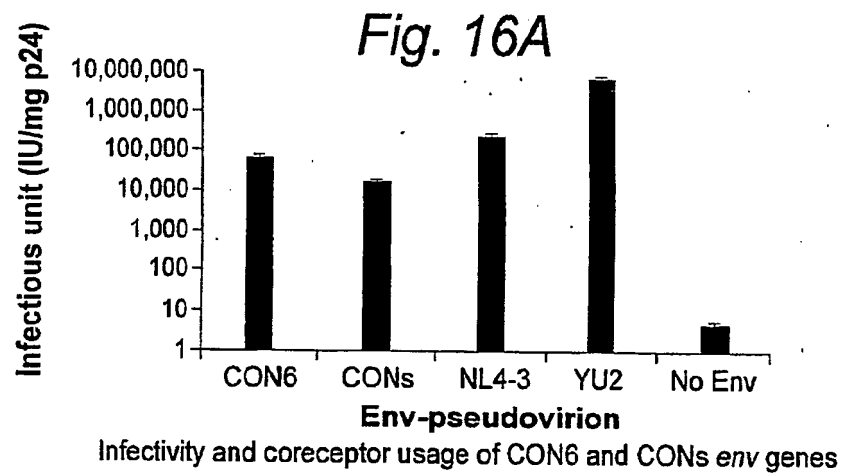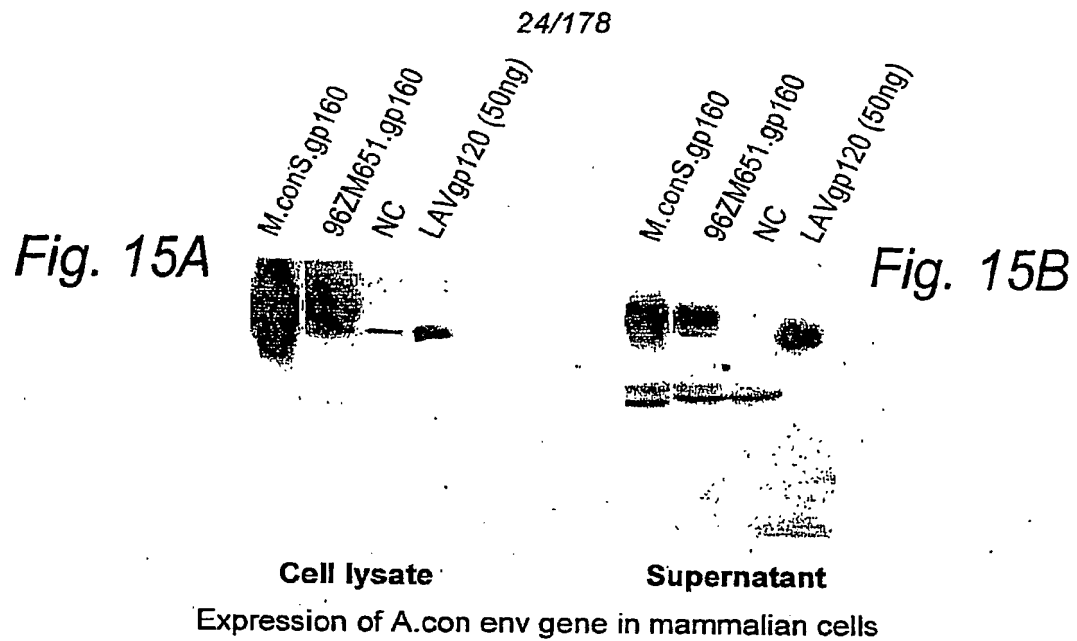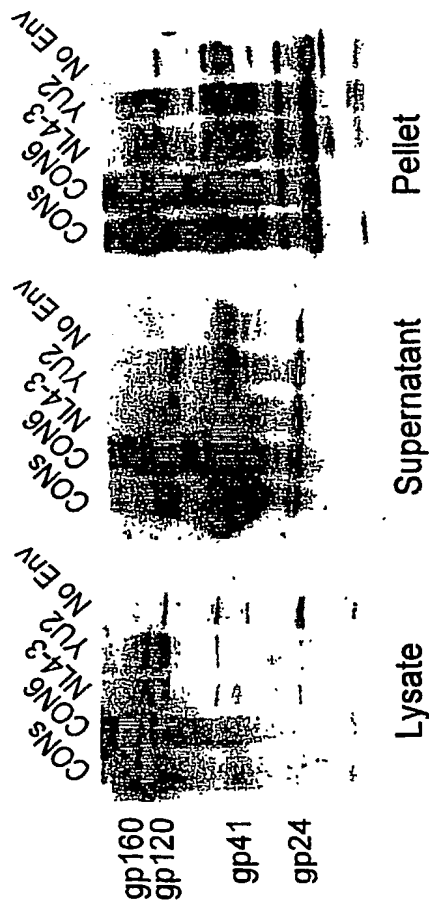CONS.gp160.4
CONS.gp160.5
CONS.gp160.6
CONS.gp160.7
CONS.gp160.8

Fig. 14B

CONs.env (gorup M consensus env gene. This one contain the
consensus sequence for variable regions in env gene. The identical
amino acid sequences as in the public domain)

```
GCCGCCGCCATGCGCGTGCGCGGCATCCAGCGCAACTGCCAGCACCTGTG
GCGCTGGGGCACCCTGATCCTGGGCATGCTGATGATCTGCTCCGCCGCCG
AGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCC
AACACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCC
AGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAAC
AACATGGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTC
CCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGAACTGCA
CCAACGTGAACGTGACCAACACCACCAACAACACCGAGGAGAAGGGCGAG
ATCAAGAACTGCTCCTTCAACATCACCACCGAGATCCGCGACAAGAAGCA
GAAGGTGTACGCCCTGTTCTACCGCCTGGACGTGGTGCCCATCGACGACA
ACAACAACAACTCCTCCAACTACCGCCTGATCAACTGCAACACCTCCGCC
ATCACCCAGGCCTGCCCCAAGGTGTCCTTCGAGCCCATCCCCATCCACTA
CTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCA
ACGGCACCGGCCCCTGCAAGAACGTGTCCACCGTGCAGTGCACCCACGGC
ATCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGCTCCCTGGCCGA
GGAGGAGATCATCATCCGCTCCGAGAACATCACCAACAACGCCAAGACCA
TCATCGTGCAGCTGAACGAGTCCGTGGAGATCAACTGCACCCGCCCCAAC
AACAACACCCGCAAGTCCATCCGCATCGGCCCCGGCCAGGCCTTCTACGC
CACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCTCCG
GCACCAAGTGGAACAAGACCCTGCAGCAGGTGGCCAAGAAGCTGCGCGAG
CACTTCAACAACAAGACCATCATCTTCAAGCCCTCCTCCGGCGGCGACCT
GGAGATCACCACCCACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGCA
ACACCTCCGGCCTGTTCAACTCCACCTGGATCGGCAACGGCACCAAGAAC
AACAACAACACCAACGACACCATCACCCTGCCCTGCCGCATCAAGCAGAT
CATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTACGCCCCCCCCATCG
AGGGCAAGATCACCTGCAAGTCCAACATCACCGGCCTGCTGCTGACCCGC
GACGGCGGCAACAACAACACCAACGAGACCGAGATCTTCCGCCCCGGCGG
CGGCGACATGCGCGACAACTGGCGCTCCGAGCTGTACAAGTACAAGGTGG
TGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAAGGCCAAGCGCCGCGTG
GTGGAGCGCGAGAAGCGCGCCGTGGGCATCGGCGCCGTGTTCCTGGGCTT
CCTGGGCGCCGCCGGCTCCACCATGGGCGCCGCCTCCATCACCCTGACCG
TGCAGGCCCGCCAGCTGCTGTCCGGCATCGTGCAGCAGCAGTCCAACCTG
CTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGTGTGGGG
CATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGG
ACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGCAAGCTGATCTGCACC
ACCACCGTGCCCTGGAACTCCTCCTGGTCCAACAAGTCCCAGGACGAGAT
CTGGGACAACATGACCTGGATGGAGTGGGAGCGCGAGATCAACAACTACA
CCGACATCATCTACTCCCTGATCGAGGAGTCCCAGAACCAGCAGGAGAAG
AACGAGCAGGAGCTGCTGGCCCTGGACAAGTGGGCCTCCCTGTGGAACTG
GTTCGACATCACCAACTGGCTGTGGTACATCAAGATCTTCATCATGATCG
TGGGCGGCCTGATCGGCCTGCGCATCGTGTTCGCCGTGCTGTCCATCGTG
AACCGCGTGCGCCAGGGCTACTCCCCCCTGTCCTTCCAGACCCTGATCCC
CAACCCCCGCGGCCCCGACCGCCCCGAGGGCATCGAGGAGGAGGGCGGCG
AGCAGGACCGCGACCGCTCCATCCGCCTGGTGAACGGCTTCCTGGCCCTG
GCCTGGGACGACCTGCGCTCCCTGTGCCTGTTCTCCTACCACCGCCTGCG
CGACTTCATCCTGATCGCCGCCCGCACCGTGGAGCTGCTGGGCCGCAAGG
GCCTGCGCCGCGGCTGGGAGGCCCTGAAGTACCTGTGGAACCTGCTGCAG
TACTGGGGCCAGGAGCTGAAGAACTCCGCCATCTCCCTGCTGGACACCAC
CGCCATCGCCGTGGCCGAGGGCACCGACCGCGTGATCGAGGTGGTGCAGC
GCGCCTGCCGCGCCATCCTGAACATCCCCCGCCGCATCCGCCAGGGCCTG
GAGCGCGCGCTGCTGTAA
```

Fig. 15A

M.conS.gp160
96ZM651.gp160
NC
LAVgp120 (50ng)

M.conS.gp160
96ZM651.gp160
NC
LAVgp120 (50ng)

Fig. 15B

Cell lysate   **Supernatant**

Expression of A.con env gene in mammalian cells

Fig. 16A

Infectivity and coreceptor usage of CON6 and CONs *env* genes

Fig. 16B

Infectivity and coreceptor usage of CON6 and CONs *env* genes

gp160
gp120

gp41

gp24

Lysate Supernatant Pellet

Env protein incorporation in CON6 and CONs Env-pseudovirions
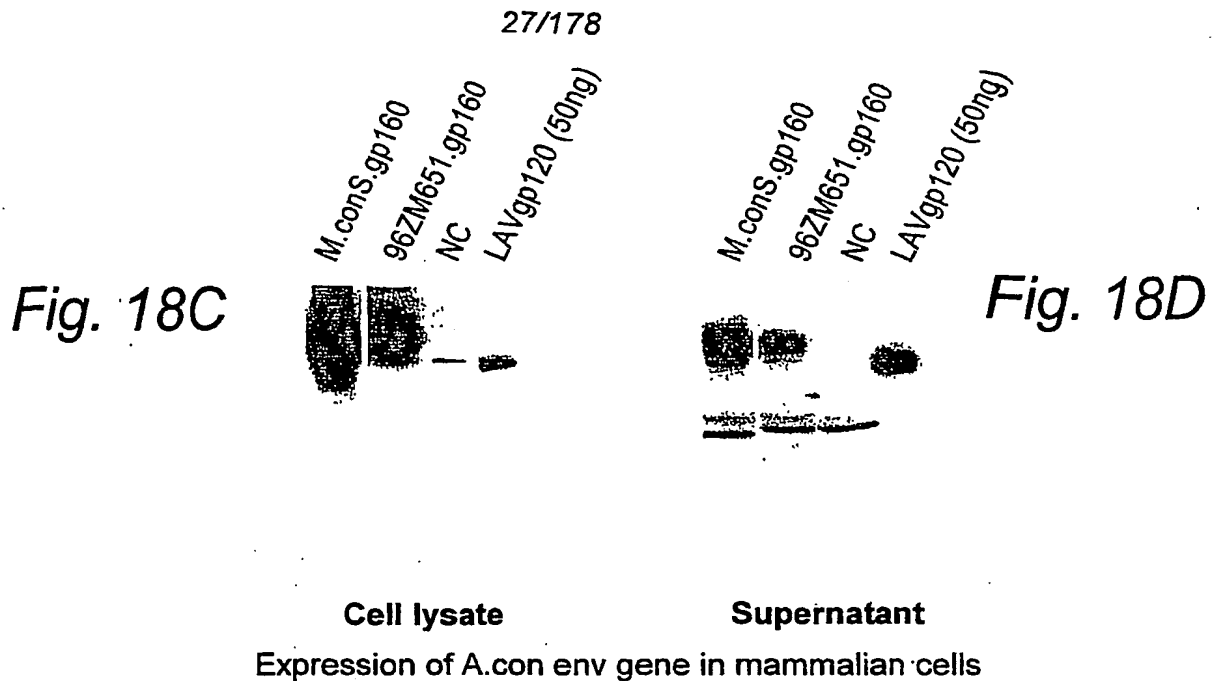
Fig. 17A   *Fig. 17B*   *Fig. 17C*

A.con.env (subtype A consensus env)
MRVMGIQRNCQHLWRWGTMLLGMIIICSAAENLWVTVYYGVPVWKDAETTLFCASDAKAYDTEVHNV
WATHACVPTDPNPQEINLENVTEEFNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLNCSNVNVTT
NITNITDNMKGEIKNCSFNMTTELRDKKQKVYSLFYKLDVVQINKSNSSSQYRLINCNTSAITQACPKVS
FEPIPIHYCAPAGFAILKCKDKEFNGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEVMIRSENITN
NAKNIIVQLTKPVKINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSRTEWNETLQKVAKQLR
KYFNNKTIIFTNSSGGDLEITTHSFNCGGEFFYCNTSGLFNSTWNGNGTKKKKNSTESNDTITLPCRIKQI
INMWQRVGQAMYAPPIQGVIRCESNITGLLLTRDGGDNNSKNETFRPGGGDMRDNWRSELYKYKVVKIEP
LGVAPTKAKRRVVEREKRAVGIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLRAIEAQQ
HLLKLTVWGIKQLQARVLAVERVLKDQQLLGIWGCSGKLICTTNVPWNSSWSNKSQSEIWDNMTWLQWDK
EISNYTDIIYNLIEESQNQQEKNEQDLLALDKWANLWNWFDISNWLMYIKIFIMIVGGLIGLRIVFAVLS
VINRVRQGYSPLSFQTHTPNPGGLDRPGRIEEEGEQGRDRSIRLVSGFLALAWDDLRSLCLFSYHRLRD
FILIAARTVELLGHSSLKGLRLGWEGLKYLWNLLLYWGRELKISAINLLDTIAIAVAGWTDRVIEIGQRI
CRAILNIPRRIRQGLERALL

Fig. 18A

Fig. 18B

A.con.env (subtype A consensus env. Identical amino acid
sequence to that in the public domain)

GCCGCCGCCATGCGCGTGATGGGCATCCAGCGCAACTGCCAGCACCTGTG
GCGCTGGGGCACCATGATCCTGGGCATGATCATCATCTGCTCCGCCGCCG
AGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGACGCC
GAGACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCC
AGGAGATCAACCTGGAGAACGTGACCGAGGAGTTCAACATGTGGAAGAAC
AACATGGTGGAGCAGATGCACACCGACATCATCTCCCTGTGGGACCAGTC
CCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGAACTGCT
CCAACGTGAACGTGACCACCAACATCACCAACATCACCGACAACATGAAG
GGCGAGATCAAGAACTGCTCCTTCAACATGACCACCGAGCTGCGCGACAA
GAAGCAGAAGGTGTACTCCCTGTTCTACAAGCTGGACGTGGTGCAGATCA
ACAAGTCCAACTCCTCCTCCCAGTACCGCCTGATCAACTGCAACACCTCC
GCCATCACCCAGGCCTGCCCCAAGGTGTCCTTCGAGCCCATCCCCATCCA
CTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAAGGACAAGGAGT
TCAACGGCACCGGCCCCTGCAAGAACGTGTCCACCGTGCAGTGCACCCAC
GGCATCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGCTCCCTGGC
CGAGGAGGAGGTGATGATCCGCTCCGAGAACATCACCAACAACGCCAAGA
ACATCATCGTGCAGCTGACCAAGCCCGTGAAGATCAACTGCACCCGCCCC
AACAACAACACCCGCAAGTCCATCCGCATCGGCCCCGGCCAGGCCTTCTA
CGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACGTGT
CCCGCACCGAGTGGAACGAGACCCTGCAGAAGGTGGCCAAGCAGCTGCGC
AAGTACTTCAACAACAAGACCATCATCTTCACCAACTCCTCCGGCGGCGA
CCTGGAGATCACCACCCACTCCTTCAACTGCGGCGGCGAGTTCTTCTACT
GCAACACCTCCGGCCTGTTCAACTCCACCTGGAACGGCAACGGCACCAAG
AAGAAGAACTCCACCGAGTCCAACGACACCATCACCCTGCCCTGCCGCAT
CAAGCAGATCATCAACATGTGGCAGCGCGTGGGCCAGGCCATGTACGCCC
CCCCCATCCAGGGCGTGATCCGCTGCGAGTCCAACATCACCGGCCTGCTG
CTGACCCGCGACGGCGGCGACAACAACTCCAAGAACGAGACCTTCCGCCC
CGGCGGCGGCGACATGCGCGACAACTGGCGCTCCGAGCTGTACAAGTACA
AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAAGGCCAAGCGC
CGCGTGGTGGAGCGCGAGAAGCGCGCCGTGGGCATCGGCGCCGTGTTCCT
GGGCTTCCTGGGCGCCGCCGGCTCCACCATGGGCGCCGCCTCCATCACCC
TGACCGTGCAGGCCCGCCAGCTGCTGTCCGGCATCGTGCAGCAGCAGTCC
AACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGAAGCTGACCGT
GTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACC
TGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGCAAGCTGATC
TGCACCACCAACGTGCCCTGGAACTCCTCCTGGTCCAACAAGTCCCAGTC
CGAGATCTGGGACAACATGACCTGGCTGCAGTGGGACAAGGAGATCTCCA
ACTACACCGACATCATCTACAACCTGATCGAGGAGTCCCAGAACCAGCAG
GAGAAGAACGAGCAGGACCTGCTGGCCCTGGACAAGTGGGCCAACCTGTG
GAACTGGTTCGACATCTCCAACTGGCTGTGGTACATCAAGATCTTCATCA
TGATCGTGGGCGGCCTGATCGGCCTGCGCATCGTGTTCGCCGTGCTGTCC
GTGATCAACCGCGTGCGCCAGGGCTACTCCCCCCTGTCCTTCCAGACCCA
CACCCCCAACCCCGGCGGCCTGGACCGCCCCGGCCGCATCGAGGAGGAGG
GCGGCGAGCAGGGCCGCGACCGCTCCATCCGCCTGGTGTCCGGCTTCCTG
GCCCTGGCCTGGGACGACCTGCGCTCCCTGTGCCTGTTCTCCTACCACCG
CCTGCGCGACTTCATCCTGATCGCCGCCCGCACCGTGGAGCTGCTGGGCC
ACTCCTCCCTGAAGGGCCTGCGCCTGGGCTGGGAGGGCCTGAAGTACCTG
TGGAACCTGCTGCTGTACTGGGGCCGCGAGCTGAAGATCTCCGCCATCAA
CCTGCTGGACACCATCGCCATCGCCGTGGCCGGCTGGACCGACCGCGTGA
TCGAGATCGGCCAGCGCATCTGCCGCGCCATCCTGAACATCCCCCGCCGC
ATCCGCCᴬ̶G̶G̶C̶C̶T̶G̶G̶A̶G̶T̶G̶C̶G̶G̶H̶C̶T̶G̶T̶A̶A̶ F 26)

SUBSTITUTE SHEET (RULE 26)

Fig. 18C

Fig. 18D

Cell lysate   **Supernatant**

Expression of A.con env gene in mammalian cells

Fig. 19A

M.con.gag (group M consensus gag. Identical amino acid sequence to that in the public domain)

GCCGCCGCCATGGGCGCCCGCGCCTCCGTGCTGTCCGGCGGCAAGCTGGA
CGCCTGGGAGAAGATCCGCCTGCGCCCCGGCGGCAAGAAGAAGTACCGCC
TGAAGCACCTGGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCCTGAAC
CCCGGCCTGCTGGAGACCTCCGAGGGCTGCAAGCAGATCATCGGCCAGCT
GCAGCCCGCCCTGCAGACCGGCTCCGAGGAGCTGCGCTCCCTGTACAACA
CCGTGGCCACCCTGTACTGCGTGCACCAGCGCATCGAGGTGAAGGACACC
AAGGAGGCCCTGGAGAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAA
GACCCAGCAGGCCGCCGCCGACAAGGGCAACTCCTCCAAGGTGTCCCAGA
ACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAGGCCATC
TCCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAGGCCTT
CTCCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCC
CCCAGGACCTGAACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCC
ATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCG
CCTGCACCCCGTGCACGCCGGCCCCATCCCCCCCGGCCAGATGCGCGAGC
CCCGCGGCTCCGACATCGCCGGCACCACCTCCACCCTGCAGGAGCAGATC
GCCTGGATGACCTCCAACCCCCCCATCCCCGTGGGCGAGATCTACAAGCG
CTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCCCCCGTGT
CCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTG
GACCGCTTCTTCAAGACCCTGCGCGCCGAGCAGGCCACCCAGGACGTGAA
GAACTGGATGACCGACACCCTGCTGGTGCAGAACGCCAACCCCGACTGCA
AGACCATCCTGAAGGCCCTGGGCCCCGGCGCCACCCTGGAGGAGATGATG
ACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTGCTGGC
CGAGGCCATGTCCCAGGTGACCAACGCCGCCATCATGATGCAGCGCGGCA
ACTTCAAGGGCCAGCGCCGCATCATCAAGTGCTTCAACTGCGGCAAGGAG
GGCCACATCGCCCGCAACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAA
GTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCA
ACTTCCTGGGCAAGATCTGGCCCTCCAACAAGGGCCGCCCCGGCAACTTC
CTGCAGTCCCGCCCCGAGCCCACCGCCCCCCCCGCCGAGTCCTTCGGCTT
CGGCGAGGAGATCACCCCCCTCCCCCAAGCAGGAGCCCAAGGACAAGGAGC
CCCCCCTGACCTCCCTGAAGTCCCTGTTCGGCAACGACCCCCTGTCCCAG
TAA

M.con.pol.nuc   *28/178*   *Fig. 19B*

```
GCCGCCGCCATGCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGACCAT
CAAGATCGGCGGCCAGCTGAAGGAGGCCCTGCTGGCCACGGGCGCCGACG
ACACCGTGCTGGAGGAGATCAACCTGCCCGGCAAGTGGAAGCCCAAGATG
ATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCT
GATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGGTGGGCCCCA
CCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGATCGGCTGCACC
CTGAACTTCCCCATCTCCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCC
CGGCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGA
TCAAGGCCCTGACCGAGATCTGCACCGAGATGGAGAAGGAGGGCAAGATC
TCCAAGATCGGCCCCGAGAACCCCTACAACACCCCCATCTTCGCCATCAA
GAAGAAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCGCGAGCTGA
ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCC
GCCGGCCTGAAGAAGAAGAAGTCCGTGACCGTGCTGGACGTGGGCGACGC
CTACTTCTCCGTGCCCCTGGACGAGGACTTCCGCAAGTACACCGCCTTCA
CCATCCCCTCCATCAACAACGAGACCCCCGGCATCCGCTACCAGTACAAC
GTGCTGCCCCAGGGCTGGAAGGGCTCCCCCGCCATCTTCCAGTCCTCCAT
GACCAAGATCCTGGAGCCCTTCCGCACCCAGAACCCCGAGATCGTGATCT
ACCAGTACATGGACGACCTGTACGTGGGCTCCGACCTGGAGATCGGCCAG
CACCGCGCCAAGATCGAGGAGCTGCGCGAGCACCTGCTGCGCTGGGGCTT
CACCACCCCCGACAAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGGATGG
GCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCCAGCTGCCC
GAGAAGGACTCCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCT
GAACTGGGCCTCCCAGATCTACCCCGGCATCAAGGTGAAGCAGCTGTGCA
AGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAG
GAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGAAGGAGCCCGT
GCACGGCGTGTACTACGACCCCTCCAAGGACCTGATCGCCGAGATCCAGA
AGCAGGGCCAGGACCAGTGGACCTACCAGATCTACCAGGAGCCCTTCAAG
AACCTCAAGACCGGCAAGTACGCCAAGATGCGCTCCGCCCACACCAACGA
CGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCACCGAGTCCATCG
TGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACC
```

```
TGGGAGACCTGGTGGACCGAGTACTGGCAGGCCACCTGGATTCCCGAGTG
GGAGTTCGTGAACACCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGA
AGGAGCCCATCGCCGGCGCCGAGACCTTCTACGTGGACGGCGCCGCCAAC
CGCGAGACCAAGCTGGGCAAGGCCGGCTACGTGACCGACCGCGGCCGCCA
GAAGGTGGTGTCCCTGACCGAGACCACCAACCAGAAAACCGAGCTGCAGG
CCATCCACCTGGCCCTGCAGGACTCCGGCTCCGAGGTGAACATCGTGACC
GACTCCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGTCCGA
GTCCGAGCTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGG
TGTACCTGTCCTGGGTGCCCGCCCACAAGGGCATCGGCGGCAACGAGCAG
GTGGACAAGCTGGTGTCCACCGGCATCCGCAAGGTGCTGTTCCTGGACGG
CATCGACAAGGCCCAGGAGGAGCACGAGAAGTACCACTCCAACTGGCGCG
CCATGGCCTCCGACTTCAACCTGCCCCCCATCGTGGCCAAGGAGATCGTG
GCCTCCTGCGACAAGTGCCAGCTGAAGGGCGAGGCCATGCACGGCCAGGT
GGACTGCTCCCCCGGCATCTGGCAGCTGGACTGCACCCACCTGGAGGGCA
AGATCATCCTGGTGGCCGTGCACGTGGCCTCCGGCTACATCGAGGCCGAG
GTGATCCCCGCCGAGACCGGCCAGGAGACCGCCTACTTCATCCTGAAGCT
GGCCGGCCGCTGGCCCGTGAAGGTGATCCACACCGACAACGGCTCCAACT
TCACCTCCGCCGCCGTGAAGGCCGCCTGCTGGTGGCCGGCATCCAGCAG
GAGTTCGGCATCCCCTACAACCCCCAGTCCCAGGGCGTGGTGGAGTCCAT
GAACAAGGAGCTGAAGAAGATCATCGGCCAGGTGCGCGACCAGGCCGAGC
ACCTCAAGACCGCCGTGCAGATGGCCGTGTTCATCCACAACTTCAAGCGC
AAGGGCGGCATCGGCGGCTACTCCGCCGGCGAGCGCATCATCGACATCAT
CGCCACCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCACCAAGATCC
AGAACTTCCGCGTGTACTACCGCGACTCCCGCGACCCCATCTGGAAGGGC
CCCGCCAAGCTGCTGTGGAAGGGCGAGGGCGCCGTGGTGATCCAGGACAA
CTCCGACATCAAGGTGGTGCCCCGCCGCAAGGCCAAGATCATCCGCGACT
ACGGCAAGCAGATGGCCGGCGACGACTGCGTGGCCGGCCGCCAGGACGAG
GACTAA
```

Fig. 19C

M.con.nef (group M consensus nef. Identical amino acid sequence
to that in the public domain)

```
GCCG CCGCCATGGGCG GCAAG TGGTCCAAGTCCTCCATCGTGGGCTGGCC
CGCCGTGCG CGAGCGCATCCG CCGCACCCACCCCGCCGCCGAGGGCGTGG
GCGCCGTGTCCCAGGACCTGGACAAGCACGGCGCCATCACCTCCTCCAAC
ACCGCCGCCAACAACCCCGACTGCGCCTGGCTGGAGGCCCAGGAGGAGGA
GGAGGAGGTGGGCTTCCCCGTGCGCCCCCAGGTGCCCCTGCGCCCCATGA
CCTACAAGGCCGCCCTGGACCTGTCCCACTTCCTGAAGGAGAAGGGCGGC
CTGGAGGGCCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTGTG
GGTGTACCACACCCAGGGCTACTTCCCCGACTGGCAGAACTACACCCCCG
GCCCCGGCATCCGCTACCCCCCTGACCTTCGGCTGGTGCTTCAAGCTGGTG
CCCGTGGACCCCGAGGAGGTGGAGGAGGCCAACGAGGGCGAGAACAACTC
CCTGCTGCACCCCATGTGCCAGCACGGCATGGAGGACGAGGAGCGCGAGG
TGCTGATGTGGAAGTTCGACTCCCGCCTGGCCCTGCGCCACATCGCCCGC
GAGCTGCACCCCGAGTACTACAAGGACTGCTAA
```

Fig. 19D

C.con.pol.nuc
```
GCCGCCGCCATGCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGTCCAT
CAAGGTGGGCGGCCAGATCAAGGAGGCCCTGCTGGCCACCGGCGCCGACG
ACACCGTGCTGGAGGAGATCAACCTGCCCGGCAAGTGGAAGCCCAAGATG
ATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCT
GATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGGTGGGCCCCA
CCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGCACC
CTGAACTTCCCCATCTCCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCC
CGGCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGA
TCAAGGCCCTGACCGCCATCTGCGAGGAGATGGAGAAGGAGGGCAAGATC
ACCAAGATCGGCCCCGAGAACCCCTACAACACCCCCGTGTTCGCCATCAA
GAAGAAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA
ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCC
GCCGGCCTGAAGAAGAAGAAGTCCGTGACCGTGCTGGACGTGGGCGACGC
CTACTTCTCCGTGCCCCTGGACGAGGGCTTCCGCAAGTACACCGCCTTCA
CCATCCCCTCCATCAACAACGAGACCCCCGGCATCCGCTACCAGTACAAC
GTGCTGCCCCAGGGCTGGAAGGGCTCCCCCGCCATCTTCCAGTCCTCCAT
GACCAAGATCCTGGAGCCCTTCCGCGCCCAGAACCCCGAGATCGTGATCT
ACCAGTACATGGACGACCTGTACGTGGGCTCCGACCTGGAGATCGGCCAG
CACCGCGCCAAGATCGAGGAGCTGCGCGAGCACCTGCTGAAGTGGGGCTT
CACCACCCCCGACAAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGGATGG
GCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCCAGCTGCCC
GAGAAGGACTCCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCT
GAACTGGGCCTCCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCA
AGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAG
GAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGAAGGAGCCCGT
GCACGGCGTGTACTACGACCCCTCCAAGGACCTGATCGCCGAGATCCAGA
AGCAGGGCCACGACCAGTGGACCTACCAGATCTACCAGGAGCCCTTCAAG
AACCTCAAGACCGGCAAGTACGCCAAGATGCGCACCGCCCACACCAACGA
CGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGTCCATCG
TGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACC
TGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATTCCCGAGTG
GGAGTTCGTGAACACCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGA
AGGAGCC░░░░░░░░░░░░░░░░░░░░░░░░░░░░░░CGCCGCCAAC
```

Fig. 19D (continued)

```
CGCGAGACCAAGATCGGCAAGGCCGGCTACGTGACCGACCGCGGCCGCCA
GAAGATCGTGTCCTGACCGGAGACCACCAACCAGAAAACGAGCTGCAGG
CCATCCAGCTGCCCTGCAGGACTTCCGGCTCCGAGGTGAACATCGTGACC
GACTCCCAGTACGCCCTGGCATCATCGAGCAGCAGGCCCGACAAGTCCGA
GTCCGAGCTGGTGAACCAGATCATGAGCAGCAGGGCATGGCGGCAACGAGCGCG
TGTACCTGTCCTGGGTGCCCGTGAAGGCATGGCGGCGGCAACGAGCAG
GTGGACAAGCTGGTGTGTCCTCCGGCATCGCAAGGTGCTGTTCCTGGACGG
CATCGACAAGGCCCAGGAGGAGCACGAGAAGTACCACTCCAACTGGCGCG
CCATGGCCTCCGAGTTCAACCTGCCCCCATCGTGGCCAAGGAGATCGTG
GCCTCCTGCGACAAGTGCCAGCTGACAGGGCGAGGCCATGCCCACCTGGAGGGCA
GGACTGCTCCCCCGGCATCTGGCCGTGCACGTGGCCTCCGGCTACATCGAGGCCGAG
AGATCATCCCCCGCCGCCGAGACCGGCCAGGAGACCGGCCTACTTCATCCTGAAGCT
GTGATCCCCCCGCCGCCGTGACCGGCCAGGAGACCGGCCTACTTCATCCTGAAGCT
GGCCGCCGCTGGCCCGTGAAGGCCGCCTGCTGGTGGGCCGGCATCCAGCAG
TCACCTCGCCGCCGTGAAGGCCGCCTGCTGGTGGGCCGGCATCCAGCAG
GAGTTCGGCATCCCTACAACCCCCAGTCCCAGGGCGTGGTGGAGTCCAT
GAACAAGGAGCTGAAGAAGATCATCGGCCAGTGCGACCAGCCGAGC
ACCTCAAGACCGCCGTGCAGATGGCCGTGTTCATCCACAACTTCAAGCGC
AAGGGCGGCATCGGCGGCTACTCCGCCGGCGAGCGCATCAGCGACATCAT
CGCCACCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCATCAAGATCC
AGAACTTCCGCGTGTACTACCGGCGACTCCCGCGACCCCATCTGGAAGGGC
CCCGCCAAGCTGCTGGAAGGTGGTGCCCGCCGCAAGGCCGAGGGCGCGAGGACAA
CTCCGACATCAAGGTGGTCGCCGGCCGGCCGACTGCCGTGGTGATCCAGGACACT
ACGGCAAGCAGCAGATGGCCGGCGGCCGACTGCCGTGCCGGCCGAGGACGAG
GACTAA
```

Fig. 19E

M.con.gag (group M consensus gag)

```
MGARASVLSGGKLDAWEKIRLRPGGKKKYRLKHLVWASRELERFALNPGLLETSEG CKQIIGQLQPA
LQTGSEELRSLYNTVATLYCVHQRIEVKDTKEALEKIEEEQNKSQQKTQQAAADKGNSSKVSQNYPIVQN
LQGQMVHQAISPRTLNAWVKVIEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKDTINE
EAAEWDRLHPVHAGPIPPGQMREPRGSDIAGTTSTLQEQIAWMTSNPPIPVGEIYKRWIILGLNKIVRMY
SPVSILDIRQGPKEPFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQNANPDCKTILKALGPGATLEEMM
TACQGVGGPGHKARVLAEAMSQVTNAAIMMQRGNFKGQRRIIKCFNCGKEGHIARNCRAPRKKGCWKCGK
EGHQMKDCTERQANFLGKIWPSNKGRPGNFLQSRPEPTAPPAESFGFGEEITPSPKQEPKDKEPPLTSLK
SLFGNDPLSQ
```

Fig. 19F

M.con.pol (group M consensus pol)
MPQITLWQRPLVTKIGGQLKEALLaTGADDTVLEEINLPGKWKPKMIGGIGGFIKVRQYDQILIEICGK
KAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALTEICTE
MEKEGKISKIGPENPYNTPIFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLD
VGDAYFSVPLDEDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRTQNPEIV
YQYMDDLYVGSDLEIGQHRAKIEELREHLLRWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIQLPEKD
SWTVNDIQKLVGKLNWASQIYPGIKVKQLCKLLRGAKALTDIVPLTEEAELELAENREILKEPVHGVYYD
PSKDLIAEIQKQGQGQDQWTYQIYQEPFKNLKTGKYAKMRSAHTNDVKQLTEAVQKIATESIVIWGKTPKFR
LPIQKETWETWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIAGAETFYVDGAANRETKLGKAGYVTD
RGRQKVVSLTETTNQKTELQAIHLALQDSGSEVNIVTDSQYALGIIQAQPDKSESELVNQIIEQLIKKEK
VYLSWVPAHKGIGGNEQVDKLVSTGIRKVLFLDGIDKAQEEHEKYHSNWRAMASDFNLPPIVAKEIVASC
DKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPV
KVIHTDNGSNFTSAAVKAACWWAGIQQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAV
FIHNFKRKGGIGGYSAGERIIDIIATDIQTKELQKQITKQNFRVYYRDSRDPIWKGPAKLLWKGEGAVV
IQDNSDIKVVPRRKAKIIRDYGKQMAGDDCVAGRQDED

Fig. 19G

M.con.nef (group M consensus nef)
MGGKWSKSSIVGWPAVRERIRRTHPAAEGVGAVSQDLDKHGAITSSNTAANNPDCAWLEAQEEEEVGFP
VRPQVPLRPMTYKAALDLSHFLKEKGGLEGLIYSKKRQEILDLWVYHTQGYFPDWQNYTPGPGIRYPLTF
GWCFKLVPVDPEEVEEANEGENNSLLHPMCQHGMEDEEREVLMWKFDSRLALRHIARELHPEYYKDC

Fig. 19H

C.con.pol (subtype C consensus pol)
MPQITLWQRPLVSIKVGGQIKEALLaTGADDTVLEEINLPGKWKPKMIGGIGGFIKVRQYDQILIEICGK
KAIGTVLVGPTPVNIIGRNMLTQLGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALTAICEE
MEKEGKITKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLD
VGDAYFSVPLDEGFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRAQNPEIVI
YQYMDDLYVGSDLEIGQHRAKIEELREHLLKWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIQLPEKD
SWTVNDIQKLVGKLNWASQIYPGIKVKQLCKLLRGAKALTDIVPLTEEAELELAENREILKEPVHGVYYD
PSKDLIAEIQKQGHDQWTYQIYQEPFKNLKTGKYAKMRTAHTNDVKQLTEAVQKIAMESIVIWGKTPKFR
LPIQKETWETWWTDYWQATWIPEWEFVNTPPLVKLWYQLEKEPIAGAETFYVDGAANRETKIGKAGYVTD
RGRQKIVSLTETTNQKTELQAIQLALQDSGSEVNIVTDSQYALGIIQAQPDKSESELVNQIIEQLIKKER
VYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGIDKAQEEHEKYHSNWRAMASEFNLPPIVAKEIVASC
DKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPV
KVIHTDNGSNFTSAAVKAACWWAGIQQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAV
FIHNFKRKGGIGGYSAGERIIDIIATDIQTKELQKQIIKQNFRVYYRDSRDPIWKGPAKLLWKGEGAVV
IQDNSDIKVVPRRKAKIIKDYGKQMAGADCVAGRQDED

Fig. 20A

B.con.gag (subtype B consensus gag. The amino acid sequence is
different from Los Alamos Database August 2002)

GCCGCCGCCATGGGCGCCCGCGCCTCCGTGCTGTCCGGCGGCGAGCTGGA
CCGCTGGGAGAAGATCCGCCTGCGCCCCGGCGGCAAGAAGAAGTACAAGC
TGAAGCACATCGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCGTGAAC
CCCGGCCTGCTGGAGACCTCCGAGGGCTGCCGCCAGATCCTGGGCCAGCT
GCAGCCCTCCCTGCAGACCGGCTCCGAGGAGCTGCGCTCCCTGTACAACA
CCGTGGCCACCCTGTACTGCGTGCACCAGCGCATCGAGGTGAAGGACACC
AAGGAGGCCCTGGAGAAGATCGAGGAGGAGCAGAACAAGTCCAAGAAGAA
GGCCCAGCAGGCCGCCGCCGACACCGGCAACTCCTCCCAGGTGTCCAGA
ACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAGGCCATC
TCCCCCCGCACCCTGAACGCCTGGGTGAAGGTGGTGGAGGAGAAGGCCTT
CTCCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCC
CCCAGGACCTGAACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCC
ATGCAGATGCTGAAGGAGACCATCAACGAGGAGGCCGCCGAGTGGGACCG
CCTGCACCCCGTGCACGCCGGCCCCATCGCCCCCGGCCAGATGCGCGAGC
CCCGCGGCTCCGACATCGCCGGCACCACCTCCACCCTGCAGGAGCAGATC
GGCTGGATGACCAACAACCCCCCCATCCCCGTGGGCGAGATCTACAAGCG
CTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCCCCCACCT
CCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTG
GACCGCTTCTACAAGACCCTGCGCGCCGAGCAGGCCTCCCAGGAGGTGAA
GAACTGGATGACCGAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCA
AGACCATCCTGAAGGCCCTGGGCCCCGCCGCCACCCTGGAGGAGATGATG
ACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTGCTGGC
CGAGGCCATGTCCCAGGTGACCAACTCCGCCACCATCATGATGCAGCGCG
GCAACTTCCGCAACCAGCGCAAGACCGTGAAGTGCTTCAACTGCGGCAAG
GAGGGCCACATCGCCAAGAACTGCCGCGCCCCCCGCAAGAAGGGCTGCTG
GAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGG
CCAACTTCCTGGGCAAGATCTGGCCCTCCCACAAGGGCCGCCCCGGCAAC
TTCCTGCAGTCCCGCCCCGAGCCCACCGCCCCCCCCGAGGAGTCCTTCCG
CTTCGGCGAGGAGACCACCACCCCCTCCCAGAAGCAGGAGCCCATCGACA
AGGAGCTGTACCCCCTGGCCTCCCTGCGCTCCCTGTTCGGCAACGACCCC
TCCTCCCAGTAA

Fig. 20B

B.con.env (subtype B consensus env. The amino acid sequence is
different from Los Alamos Database August 2002)

```
GCCG CCGCCATGCGCG TGAAGGG CA TCCGCAAGAA CTACCAG CAC CTGTG
GCGCTGGGG CA CCATG CTGCTGGGCATG CT GATGA TCTGC TCCGC CGCCG
AGAAG CT GTGGG TGACCGTGTA CTACGG CG TGCCCGTGTG GAAGGA GGCC
ACCA CCACC CTGTTCT GCGCCTC CGACG CCAAGGC CTACGA CACC GAGGT
GCACAA CGTGTGGG CCACCCA CGCCTGCGTGCCCA CCGAC CC CAA CCCCC
AGGA GGTGG TG CTGGA GAACG TGAC CGA GAA CTTCAA CAT GTGGAA GAAC
AACA TGGTGGA GCAGA TGCACGAGGACA TCATCTC CCTGTGGGAC CAGT C
CCTGAAGC CCTGCGTGAAGC TGACC CC CCT GTGCG TGACC CTGAA CTGCA
CCGA CCTGAAGAACAA CCTGC TGAA CAC CAACTC CTCCTC CGGCGAGAAG
ATGGAGAAGGG CGAGA TCAAGAA CT GCT CC TTCAA CATCA CCACC TC CAT
CCGCGA CAAGG TGCAG AAGGA GTACGCC CTGTT CTACAAG CTGGA CGTGG
TGCC CATCGA CAACAA CAACAA CAC CTC CTACCGC CTGAT CT CCTGCAAC
AC CT CCG TGATCAC CCAGG CCTGC CC CAAGG TGTC CTTCGAG CCCATC CC C
CATC CA CTA CT GCGCC CC CGC CGG CTTCGC CATC CT GAAG TGCAA CGACA
AGAA GTT CAAC GGCAC CGGCC CCTG CAC CAA CGTGTC CAC CGTGCAG TG C
ACCCA CGGCAT CCGCC CCGTGG TGT CCA CC CAGCT GCTGC TGAA CGG CT C
CCTGG CCGAGG AGGAG GTGGTGA TC CGC TC CGA GAA CTTCAC CGA CAAC G
C CAAGACCAT CATCGTGCAG CTGAA CGAGTCCG TG GAGAT CAACT GCACC
CGCC CCAACAA CAACA CCCGCAAGT CCA TC CACAT CGGCC CCGGC CGCGC
CTTC TA CAC CA CCGGC GAGAT CA TCGGC GA CATCCGCCAG GCCCA CTGCA
ACAT CT CCC GCGCCAA GTGGAACAA CAC CC TGAAG CAGAT CGTGAAGAAG
CTGCGCGAGCA GTTCGG CAACAA GACCATCGTGTT CAACCAGTCCTCCGG
CGGCGA CCC CGAGATC GTGATGCACT CC TT CAACTGCGGCGG CGAGTT CT
TCTA CTGCAACACCAC CCAGC TGTT CAA CT CCACCTGGAA CGACAACGGC
ACCTGGAACAA CAC CAAGGACAAGAACA CCATCAC CCTGC CCTGC CGCAT
CAAG CAGATCA TCAACAT GTGGCAG GAGGT GGGCAAGGCCATGTA CGCCC
CCCC CATCCGCGGCCA GATCCGCTGCTC CT CCAACATCAC CGGCCTGCTG
CTGAC CCGCGAC GGCGG CAACAA CAACAA CGACAC CGAGA TC TTC CGC CC C
CGGCGG CGG CGA CATG CGCGA CAAC TGGCG CTCCGAGCTG TA CAAGTACA
AGGT GG TGAAGATCGA GCCCC TGGG CGTGG CCCCCACCAAGG CCAAG CGC
CGCG TGGTG CA GCGCGAGAAG CG CG CCG TGGGCAT CGGCG CCATGTT CCT
GGGCTT CCTGG GCGCCGCCGG CT CCACCATGGGCG CCGCC TC CATGA CCC
TGAC CG TG CAGG C CCG CCAG CTG CT GTC CGGCATCGTGCA GCAGCAGAA C
AACCTGC TGCG CGCCA TCGAGGC CCAGCAG CACCTGCTGCAG CTGACCGT
GTGGGG CAT CAAGCAG CTGCAGG C C CGCGTGCTGG CCGTGGAGCG CTAC C
TGAAGGACCAG CAGCT GCTGGGCAT CTGGGGCTGCTCCGG CAAGCTGATC
TGCA CCAC CAC CGTGCC CTGGAA CG CCT CCTGGTC CAACAAG TCC CTGGA
CGAGATCTGGGA CAACATGACCTGGATGGAGTGGGAGCGC GAGAT CGACA
ACTA CA CCT CC CTGAT CTACA CC CT GAT CGAGGAG TCCCAGAACCAG CAG
GAGAAGAACGA GCAGGAG CTG CTGGAGC TG GACAAGTGGG CCTCCCTGTG
GAACTGGTT CGACATCACCAACT GGCTGTG.GTACA TCAAGAT CTT CATCA
TGATCGTGGGCGGCCTGATCGGC CTGCGCA TCGTGTTCGCCGTGCTGTCC
ATCGTGAACCGCGTGCGCCAGGGCTACTCC CC CCTGTCCTTC CAGACCCG
CCTGCC CGC CC CCCGCGGCCCCGACCGC CC CGAGGGCATCGA GGAGGAGG
GCGG CGAGCGCGACCG CGACCGCTC CGGCCGCCTGGTGGA CGGCTTC CTG
GCCCTGATCTGGGACGACCTGCGCT CCCTGTGCCTGTTCT CCTACCACCG
CCTGCG CGA CCTGCTG CTGATCG TGACCCG CATCG TGGAG CTGCTGGGCC
GC CGCGG CT GGGAGGT GCTGAAGTACTG GTGGAACCTGCTGCAGTACTGG
TCCCAGGAG CTGAAGAACTCCGCCGTGTCCCTGCTGAACG CCACCGCCAT
CGCCGTGGCCGAGGGCACCGACCGCGTGAT CGAGGTGGTGCAGCGCGCCT
GCCG CGCCA TC CTGCA CATCCCC CG CCG CA TCCGC CAGGG CCTGGAGCGC
GCCCTGCTGTAA
```

Fig. 20B

B.con.env (subtype B consensus env. The amino acid sequence is different from Los Alamos Database August 2002)

```
GCCGCCGCCATGCGCGTGAAGGGCATCCGCAAGAACTACCAGCACCTGTG
GCGCTGGGGCACCATGCTGCTGGGCATGCTGATGATCTGCTCCGCCGCCG
AGAAGCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCC
ACCACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGACACCGAGGT
GCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCC
AGGAGGTGGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAAC
AACATGGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTC
CCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGAACTGCA
CCGACCTGAAGAACAACCTGCTGAACACCAACTCCTCCTCCGGCGAGAAG
ATGGAGAAGGGCGAGATCAAGAACTGCTCCTTCAACATCACCACCTCCAT
CCGCGACAAGGTGCAGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGG
TGCCCATCGACAACAACAACAACACCTCCTACCGCCTGATCTCCTGCAAC
ACCTCCGTGATCACCCAGGCCTGCCCCAAGGTGTCCTTCGAGCCCATCCC
CATCCACTACTGCGCCCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACA
AGAAGTTCAACGGCACCGGCCCCTGCACCAACGTGTCCACCGTGCAGTGC
ACCCACGGCATCCGCCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGCTC
CCTGGCCGAGGAGGAGGTGGTGATCCGCTCCGAGAACTTCACCGACAACG
CCAAGACCATCATCGTGCAGCTGAACGAGTCCGTGGAGATCAACTGCACC
CGCCCCAACAACAACACCCGCAAGTCCATCCACATCGGCCCCGGCCGCGC
CTTCTACACCACCGGCGAGATCATCGGCGACATCCGCCAGGCCCACTGCA
ACATCTCCCGCGCCAAGTGGAACAACACCCTGAAGCAGATCGTGAAGAAG
CTGCGCGAGCAGTTCGGCAACAAGACCATCGTGTTCAACCAGTCCTCCGG
CGGCGACCCCGAGATCGTGATGCACTCCTTCAACTGCGGCGGCGAGTTCT
TCTACTGCAACACCACCCAGCTGTTCAACTCCACCTGGAACGACAACGGC
ACCTGGAACAACACCAAGGACAAGAACACCATCACCCTGCCCTGCCGCAT
CAAGCAGATCATCAACATGTGGCAGGAGGTGGGCAAGGCCATGTACGCCC
CCCCCATCCGCGGCCAGATCCGCTGCTCCTCCAACATCACCGGCCTGCTG
CTGACCCGCGACGGCGGCAACAACAACAACGACACCGAGATCTTCCGCCC
CGGCGGCGGCGACATGCGCGACAACTGGCGCTCCGAGCTGTACAAGTACA
AGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAAGGCCAAGCGC
CGCGTGGTGCAGCGCGAGAAGCGCGCCGTGGGCATCGGCGCCATGTTCCT
GGGCTTCCTGGGCGCCGCCGGCTCCACCATGGGCGCCGCCTCCATGACCC
TGACCGTGCAGGCCCGCCAGCTGCTGTCCGGCATCGTGCAGCAGCAGAAC
AACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAGCTGACCGT
GTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACC
TGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGCAAGCTGATC
TGCACCACCACCGTGCCCTGGAACGCCTCCTGGTCCAACAAGTCCCTGGA
CGAGATCTGGGACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGACA
ACTACACCTCCCTGATCTACACCCTGATCGAGGAGTCCCAGAACCAGCAG
GAGAAGAACGAGCAGGAGCTGCTGGAGCTGGACAAGTGGGCCTCCCTGTG
GAACTGGTTCGACATCACCAACTGGCTGTGGTACATCAAGATCTTCATCA
TGATCGTGGGCGGCCTGATCGGCCTGCGCATCGTGTTCGCCGTGCTGTCC
ATCGTGAACCGCGTGCGCCAGGGCTACTCCCCCCTGTCCTTCCAGACCCG
CCTGCCCGCCCCCCGCGGCCCCGACCGCCCCGAGGGCATCGAGGAGGAGG
GCGGCGAGCGCGACCGCGACCGCTCCGGCCGCCTGGTGGACGGCTTCCTG
GCCCTGATCTGGGACGACCTGCGCTCCCTGTGCCTGTTCTCCTACCACCG
CCTGCGCGACCTGCTGCTGATCGTGACCCGCATCGTGGAGCTGCTGGGCC
GCCGCGGCTGGGAGGTGCTGAAGTACTGGTGGAACCTGCTGCAGTACTGG
TCCCAGGAGCTGAAGAACTCCGCCGTGTCCCTGCTGAACGCCACCGCCAT
CGCCGTGGCCGAGGGCACCGACCGCGTGATCGAGGTGGTGCAGCGCGCCT
GCCGCGCCATCCTGCACATCCCCCGCCGCATCCGCCAGGGCCTGGAGCGC
GCCCTGCTGTAA
```

Fig. 20C

B.con.gag (subtype B consensus gag)

MGARASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERFAVNPGLLETSEGCRQILGQLQPSLQT
GSEELRSLYNTVATLYCVHQRIEVKDTKEALEKIEEEQNKSKKKAQQAAADTGNSSQVSQNYPIVQNLQG
QMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAA
EWDRLHPVHAGPIAPGQMREPRGSDIAGTTSTLQEIGWMTNNPPIPVGEIYKRWIILGLNKIV RMYSPT
SILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQNANPDCKTILKALGPAATLEEMMTAC
QGVGGPGHKARVLAEAMSQVTNSATIMMQRGNFRNQRKTVKCFNCGKEGHIAKNCRAPRKKGCWKCGKEG
HQMKDCTERQANFLGKIWPSHKGRPGNFLQSRPEPTAPPEESFRFGEETTTPSQKQEPIDKELYPLASLR
SLFGNDPSSQ

Fig. 20D

B.con.env (subtype B consensus env)

MRVKGIRKNYQHLWRWGTMLLGMLMICSAAEKLWVTV YYGVPVWKEATTTLFCASDAKAYDTEVHNVWAT
HACVPTDPNPQEVVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLKNNLLNT
NSSSGEKMEKGEIKNCSFNITTSIRDKVQKEYALFYKLDVVPIDNNNNTSYRLISCNTSVITQACPKVSF
EPIPIHYCAPAGFAILKCNDKKFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAEEEVVIRSENFTDN
AKTIIVQLNESVEINCTRPNNNTRKSIHIGPGRAFYTTGEIIGDIRQAHCNISRAKWNNTLKQIVKKLRE
QFGNKTIVFNQSSGGDPEIVMHSFNCGGEFFYCNTTQLFNSTWNDNGTWNNTKDKNTITLPCRIKQIINM
WQEVGKAMYAPPIRGQIRCSSNITGLLLTRDGGNNNNDTEIFRPGGGDMRDNWRSELYKYKVKIEPLGV
APTKAKRRVVQREKRAVGIGAMFLGFLGAAGSTMGAASMLTVQARQLLSGIVQQNNLLRAIEAQQHLL
QLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTTVPWNASWSNKSLDEIWDNMTWMEWEREID
NYTSLIYTLIEESQNQQEKNEQELLELDKWASLWNWFDITNWLWYIKIFIMIVGGLIGLRIVFAVLSIVN
RVRQGYSPLSFQTRLPAPRGPDRPEGIEEEGGERDRDRSGRLVDGFLALIWDDLRSLCLFSYHRLRDLLL
IVTRIVELLGRRGWEVLKYWWNLLQYWSQELKNSAVSLLNATAIAVAEGTDRVIEVVQRACRAILHIPRR
IRQGLERALL

Fig. 21

Expression of subtype B consensus *env* and *gag* genes in 293T cells. Plasmids containing codon-optimized subtype B consensus *gp160, gp140,* and *gag* genes were transfected into 293T cells, and protein expression was examined by Western Blot analysis of cell lysates. 48-hours post-transfection, cell lysates were collected, total protein content determined by the BCA protein assay, and 2 µg of total protein was loaded per lane on a 4-20% SDS-PAGE gel. Proteins were transferred to a PVDF membrane and probed with serum from an HIV-1 subtype B infected individual.

Fig. 22

Co-receptor usage of subtype B consensus envelopes.
Pseudotyped particles containing the subtype B consensus gp160 Env
were incubated with DEAE-Dextran treated JC53-BL cells in the
presence of AMD3100 (a specific inhibitor of CXCR4), TAK779 (a
specific inhibitor of CCR5), and AMD3000+TAK779 to determine co-
receptor usage. NL4.3, an isolate known to utilize CXCR4 and YU-2,
a known CCR5-using isolate, were included as controls.

Fig. 23A

Trans complementation of *env*-deficient HIV-1 with codon-optimized subtype B consensus *gp160* and *gp140* genes.

Plasmids containing codon-optimized, subtype B consensus *gp160* or *gp140* genes were co-transfected into 293T cells with an HIV-1/SG3Δenv provirus. 48-hours post-transfection cell supernatants containing pseudotyped virus were harvested, clarified in a tabletop centrifuge, filtered through a 0.2µM filter, and pellet through a 20% sucrose cushion. Quantification of p24 in each virus pellet was determined using the Coulter HIV-1 p24 antigen assay; 25ng of p24 was loaded per lane on a 4-20% SDS-PAGE gel. Proteins were transferred to a PVDF membrane and probed with anti-HIV-1 antibodies from infected HIV-1 subtype B patient serum. *Trans* complementation with a rev-dependent NL4.3*env* was included for control.

Fig. 23B
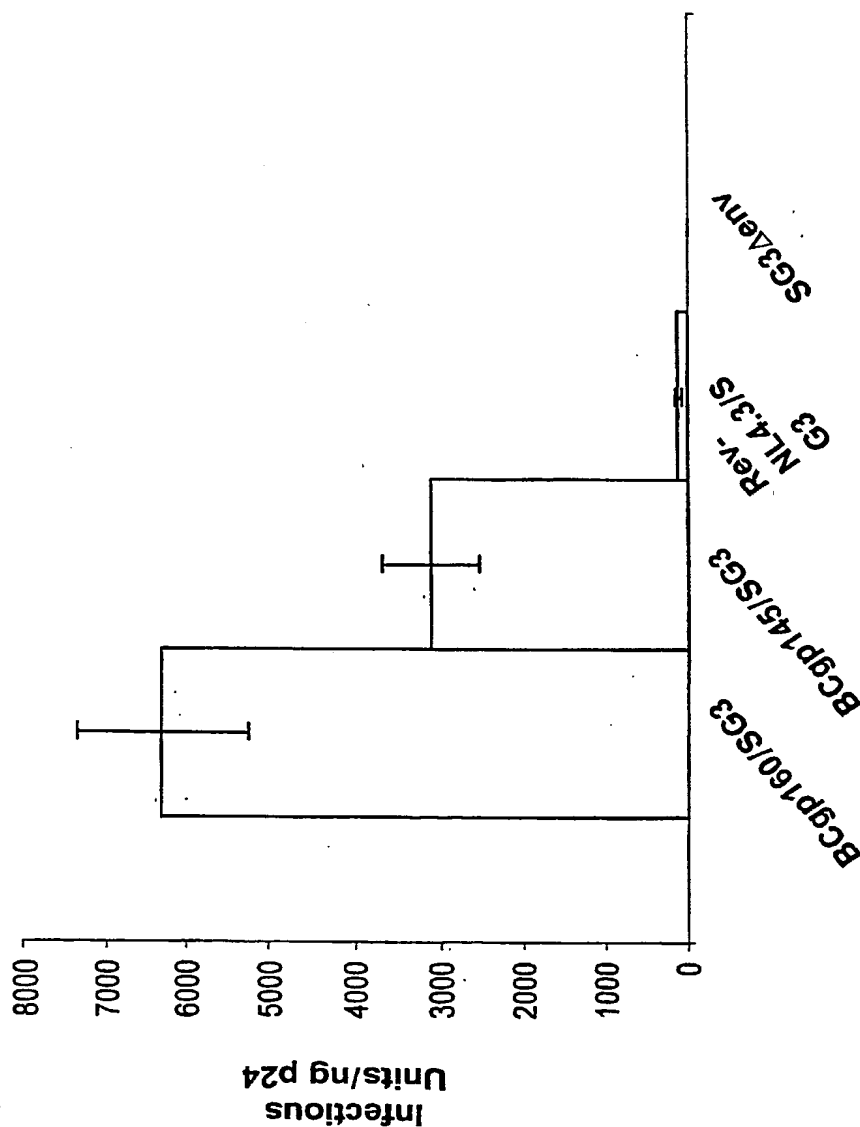
Infectivity of virus particles containing the subtype B concensus envelope.
Infectivity of pseudotyped virus containing consensus B gp160 or gp140 was determined using the JC53-BL assay. Sucrose cushion purified virus particles were assayed by the Coulter p24 antigen assay, and 5-fold serial dilutions of each pellet were incubated with DEAE-Dextran treated JC53-BL cells. Following a 48-hour incubation period, cells were fixed and stained to visualize β-galactosidase expressing cells. Infectivity is expressed as infectious units per ng of p24.

Fig. 24B
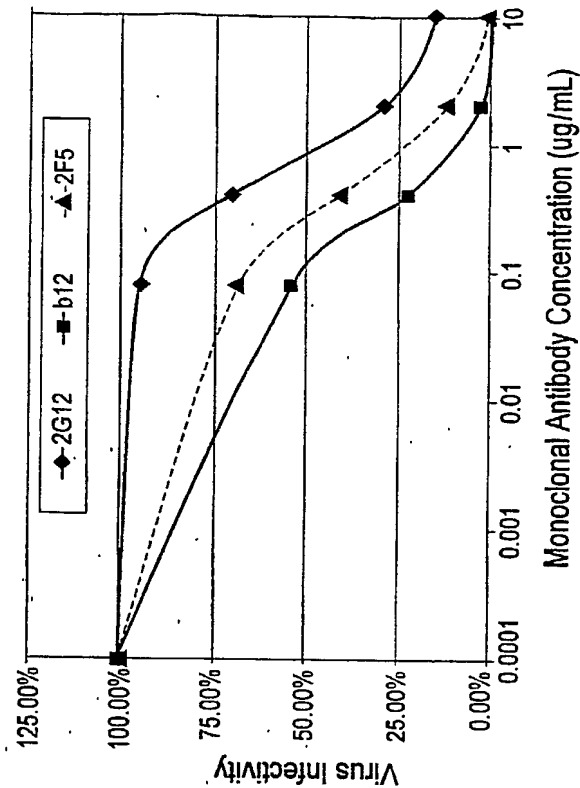
Neutralization of Pseudovirions containing
NL4.3 Env (gp160)

Fig. 24A
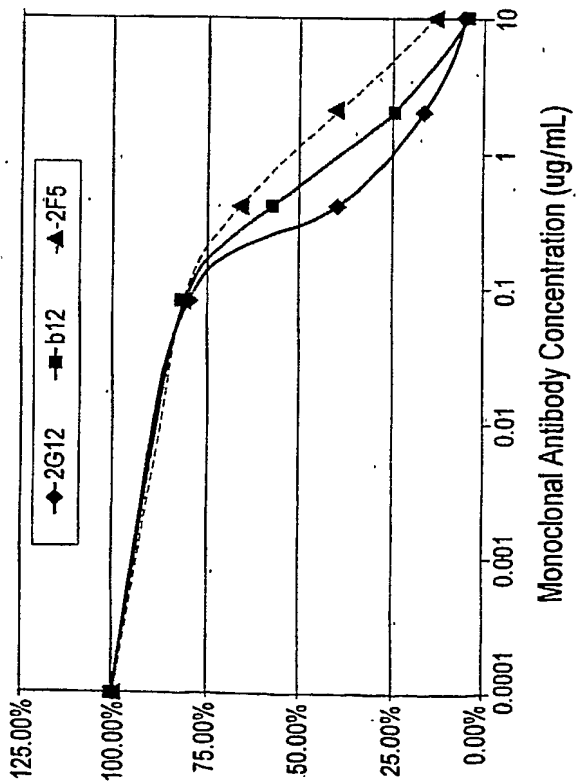
Neutralization of Pseudovirions containing Subtype B
consensus Env (gp160)
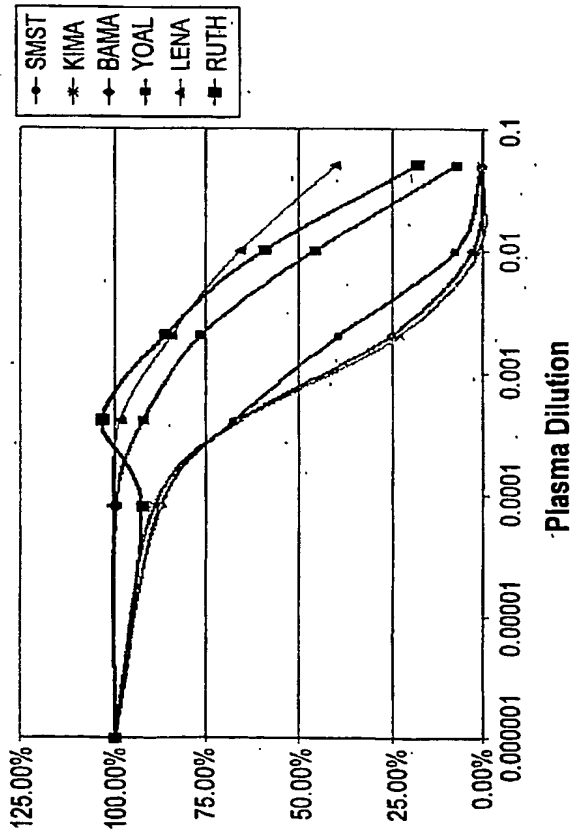
Fig. 24D

Neutralization of Pseudovirions
containing NL4.3 Env (gp160)

Fig. 24C
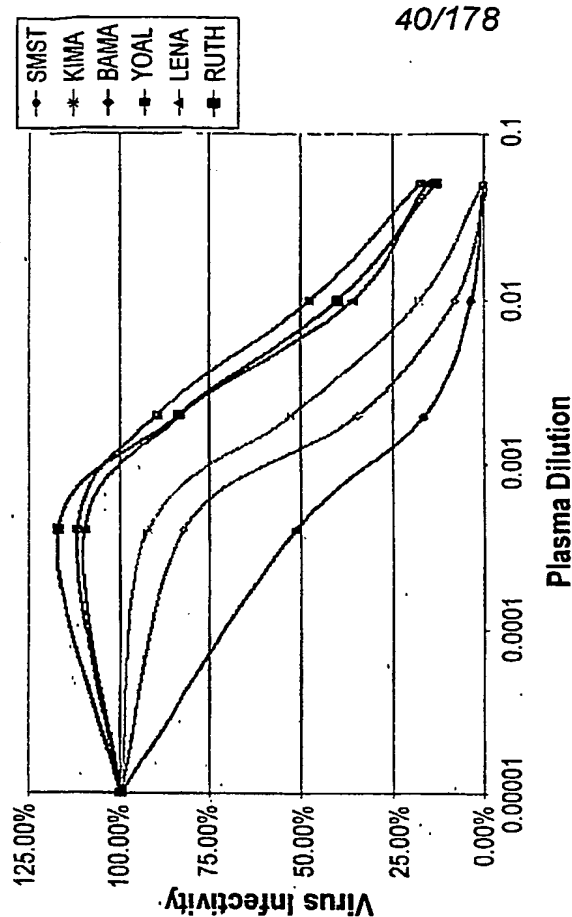
Neutralization of Pseudovirions containing
Subtype B consensus Env (gp160)

Neutralization sensitivity of virions containing subtype B concensus gp160 envelope.

Equivalent amounts of pseudovirions containing the subtype B consensus or NL4.3 Env (gp160) (1,500 infectious units) were preincubated with three different monoclonal neutralizing antibodies and a panel of plasma samples from HIV-1 subtype B infected individuals, and then added to the JC53-BL cell monolayer in 96-well plates. Plates were cultured for two days and luciferase activity was measured as an indicator of viral infectivity. Virus infectivity was calculated by dividing the luciferase units (LU) produced at each concentration of antibody by the LU produced by the control infection. The mean 50% inhibitory concentration (IC$_{50}$) and the actual % neutralization at each antibody dilution were then calculated for each virus. The results of all luciferase experiments were confirmed by direct counting of blue foci in parallel infections.

Fig. 25A

Density and p24 analysis of sucrose gradient fractions.

0.5ml fractions were collected from a 20-60% sucrose gradient. Fraction number 1 represents the most dense fraction taken from the bottom of the gradient tube. Density was measured with a refractometer and the amount of p24 in each fraction was determined by the Coulter p24 antigen assay. Fractions 6-9, 10-15, 16-21, and 22-25 were pooled together and analyzed by Western Blot. As expected, virions sedimented at a density of 1.16-1.18 g/ml.

Fig. 25B
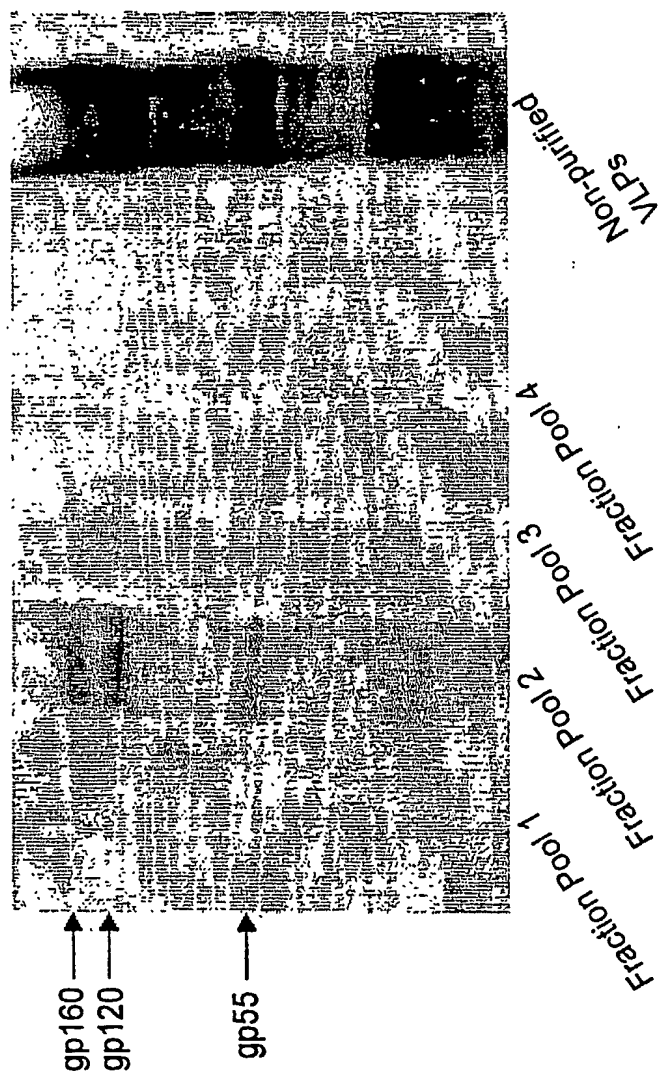
VLP production by co-transfection of subtype B consensus *gag* and *env* genes.

293T cells were co-transfected with subtype B consensus *gag* and *env* genes. Cell supernatants were harvested 48-hours post-transfection, clarified through at 20% sucrose cushion, and further purified through a 20-60% sucrose gradient. Select fractions from the gradient were pooled, added to 20ml of PBS, and centrifuged overnight at 100,000 x g. Resuspended pellets were loaded onto a 4-20% SDS-PAGE gel, proteins were transferred to a PVDF membrane, and probed with plasma from an HIV-1 subtype B infected individual.

Fig. 26A

Year 2000 Con-S 140CFI.Env

MRVRGIQRNCQHLWRWGTLILGMLMICSAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDTEVH
NVWATHACVPTDPNPQEIVLENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNC
TNVNVTNTTNNTEEKGEIKNCSFNITTEIRDKKQKVYALFYRLDVVPIDDNNNNSSNYRLINCNT
SAITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQCTHGIKPVVSTQLLLNG
SLAEEEIIIRSENITNNAKTIIVQLNESVEINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQA
HCNISGTKWNKTLQQVAKKLREHFNNKTIIFKPSSGGDLEITTHSFNCRGEFFYCNTSGLFNSTW
IGNGTKNNNNTNDTITLPCRIKQIINMWQGVGQAMYAPPIEGKITCKSNITGLLLTRDGGNNNTN
ETEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKAKLTVQARQLLSGIVQQQSNLLRAIEAQ
QHLLQLTVWGIKQLQARVLAVERYLKDQQLEIWDNMTWMEWEREINNYTDIIYSLIEESQNQQEK
NEQELLALDKWASLWNWFDITNWLW

A gp140 CFI is referred to HIV-1 envelope design with the cleavage-site-deleted (C), fusion-site-deleted
(F) and gp41 immunodominant region-deleted (I) in addition to the deletion of transmembrane and
cytoplasmic domains.

Fig. 26B

Codon-optimized Year 2000 Con-S 140CFI. seq

ATGCGCGTGCGCGGCATCCAGCGCAACTGCCAGCACCTGTGGCGCTGGGGCACCCTGATCCTGGG
CATGCTGATGATCTGCTCCGCCGCCGAGAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGT
GGAAGGAGGCCAACACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGACACCGAGGTGCAC
AACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCTGGAGAA
CGTGACCGAGAACTTCAACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCATCT
CCCTGTGGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGAACTGC
ACCAACGTGAACGTGACCAACACCACCAACAACACCGAGGAGAAGGGCGAGATCAAGAACTGCTC
CTTCAACATCACCACCGAGATCCGCGACAAGAAGCAGAAGGTGTACGCCCTGTTCTACCGCCTGG
ACGTGGTGCCCATCGACGACAACAACAACAACTCCTCCAACTACCGCCTGATCAACTGCAACACC
TCCGCCATCACCCAGGCCTGCCCCAAGGTGTCCTTCGAGCCCATCCCCATCCACTACTGCGCCCC
CGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACCGGCCCCTGCAAGAACG
TGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGC
CGTGCAGCTGAACGAGTCCGTGGAGATCAACTGCACCCGCCCCAACAACAACACCCGCAAGTCCA
TCCGCATCGGCCCCGGCCAGGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCC
CACTGCAACATCTCCGGCACCAAGTGGAACAAGACCCTGCAGCAGGTGGCCAAGAAGCTGCGCGA
GCACTTCAACAACAAGACCATCATCTTCAAGCCCTCCTCCGGCGGCGACCTGGAGATCACCACCC
ACTCCTTCAACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCGGCCTGTTCAACTCCACCTGG
ATCGGCAACGGCACCAAGAACAACAACAACACCAACGACACCATCACCCTGCCCTGCCGCATCAA
GCAGATCATCAACATGTGGCAGGGCGTGGGCCAGGCCATGTACGCCCCCCCCATCGAGGGCAAGA
TCACCTGCAAGTCCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAACAACAACACCAAC
GAGACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAACTGGCGCTCCGAGCTGTACAA
GTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAAGGCCAAGCTTACCGTGCAGG
CCCGCCAGCTGCTGTCCGGCATCGTGCAGCAGCAGTCCAACCTGCTGCGCGCCATCGAGGCCCAG
CAGCACCTGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGA
GCGCTACCTGAAGGACCAGCAGCTCGAGATCTGGGACAACATGACCTGGATGGAGTGGGAGCGCG
AGATCAACAACTACACCGACATCATCTACTCCCTGATCGAGGAGTCCCAGAACCAGCAGGAGAAG
AACGAGCAGGAGCTGCTGGCCCTGGACAAGTGGGCCTCCCTGTGGAACTGGTTCGACATCACCAA
CTGGCTGTGGTGAGGATCC

Fig. 27

Individual C56BL/6 Mouse T Cell Responses to HIV-1 Envelope Peptides
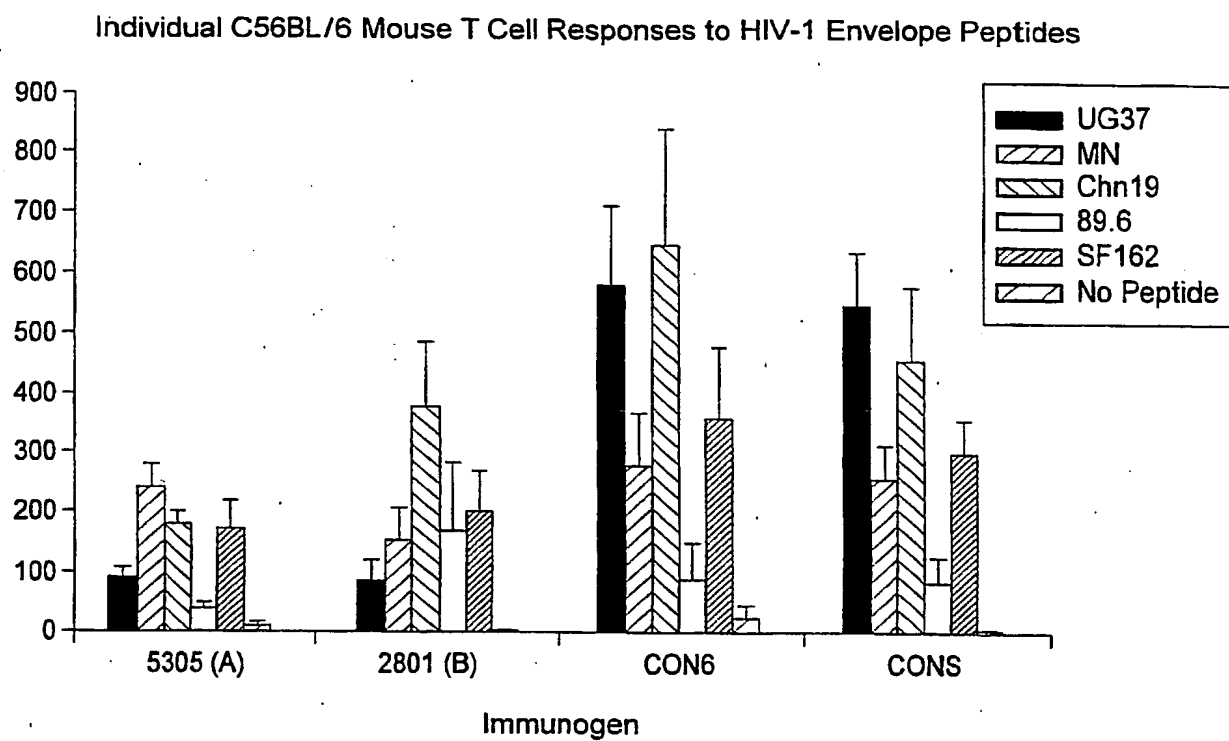
Fig. 28A

Design of expression-optimized HIV-1 envelope gp140CF

Con-B-2003 Env.pep (841 a.a.)*

MRVKGIRKNYQHLWRWGTMLLGMLMICSAAEKLWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEVVL
ENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLMNATNTNTTIIYRWRGEIKNCSFNITTSIRDKVQKEY
ALFYKLDVVPIDNDNTSYRLISCNTSVITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCTNVSTVQCTHGIRPVVSTQ
LLLNGSLAEEEVVIRSENFTDNAKTIIVQLNESVEINCTRPNNNTRKSIHIGPGRAFYTTGEIIGDIRQAHCNISRAKWNNTLKQ
IVKKLREQFGNKTIVFNQSSGGDPEIVMHSFNCGGEFFYCNTTQLFNSTWNGTWNNTEGNITLPCRIKQIINMWQEVGKAMYAPP
IRGQIRCSSNITGLLLTRDGGNNETEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRAVGIGAMFLGFLGA
AGSTMGAASMTLTVQARQLLSGIVQQQNNLLRAIEAQQHLLQLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTAVPW
NASWSNKSLDEIWDNMTWMEWEREIDNYTSLIYTLIEESQNQQEKNEQELLELDKWASLWNWFDITNWLWYIKIFIMIVGGLVGL
RIVFAVLSIVNRVRQGYSPLSFQTRLPAPRGPDRPEGIEEGGERDRSGRLVDGFLALIWDDLRSLCLFSYHRLRDLLLIVTR

*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF design and the "W" underlined with red color is the last amino acid at the C terminus, and all the remaining amino acids after the "W" will be deleted in 140CF design.

Fig. 28B

Con-B-140CF.pep (632 a.a.)
Nick name: 002

MRVKGIRKNYQHLWRWGTMLLGMLMICSAAEKLWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEVVL
ENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLMNATNTNTTIIYRWRGEIKNCSFNITTSIRDKVQKEY
ALFYKLDVVPIDNDNTSYRLISCNTSVITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCTNVSTVQCTHGIRPVVSTQ
LLLNGSLAEEEVVIRSENFTDNAKTIIVQLNESVEINCTRPNNNTRKSIHIGPGRAFYTTGEIIGDIRQAHCNISRAKWNNTLKQ
IVKKLREQFGNKTIVFNQSSGGDPEIVMHSFNCGGEFFYCNTTQLFNSTWNGTWNNTEGNITLPCRIKQIINMWQEVGKAMYAPP
IRGQIRCSSNITGLLLTRDGGNNETEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKAKTLTVQARQLLSGIVQQQNNLLRA
IEAQQHLLQLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTAVPWNASWSNKSLDEIWDNMTWMEWEREIDNYTSLIY
TLIEESQNQQEKNEQELLELDKWASLWNWFDITNWLW*

*Amino acids seen in blue color is for easy identification of the junction of the deleted fusion cleavage site.

Fig. 28C

Codon-opitmized Con-B 140CF.seq (1927 nt.)

Nick name: 002

```
TTCAGTCGACGGCCACCATGAGGGTGAAGGGTATTCGGAAAAAATTACCAACACCTGTGGCGCTGGGGAACCATGCTCCTTGGTAT
GCTGATGATTGCAGTCGCCGCGAGAAGAACTTGGGTAACTGTGTACTACGGCGTTCCTGTCTGGAAGGAAGCTACAACCACTCTT
TTTTGTGCATCCGACGGCTAAAGCTTACGACACAGAAGTGCATAATGTTTGGGCCACCCATGCTTGCGTCCCTACAGATCCCAACC
CCCAGGAAGTCGTTCCTTGAGAATGTCACAGAGAATTTTAACATGTGGAAGAATAATATGGTAGACAACAAATGCACGAAGACATTAT
TAGCCTGTGGGACCAGTCCTTGAGCCCTGCGTGAAACTCACTCCACTTTGCGTCACACTTAACTGTACTGATTTGATGAACGCA
ACCAACACAAATACTACTATTATATATCGCGTGAGGGGAAATCAAGAAACTGCTCTTTCAACATCACCACTTCCATAAGGGATA
AGGTCCAGAAAGAATATGCCCTGTTTTATAAACTTGATGTGGTCCCAAAGGTCAGCTTTGAACCAATACCCATTCCACCGTGCC
TAACACATCCGTGATTACCCAAGCTTGCCCAAAGGTCACCGGACCTTGCCGGAAGAGAACCTTGCCAATGTCTTCACGGAATCAGAC
ATCCTCAAGTGTAACGACAAAAAAAATTCAATGGGACGGGTCTTCGCGGAAGAGAACGGTCGTGATTAGAAGCGAAAAACTTTACCGATAACGC
CTGTTGTCAGTACCCAACTCCTCTTGAACGGTCCATTTTATACAACTGGCAGAAGTCATTGTGACATCGAAAAATCTATTCACATA
GGGCCCGGCCCGCGCATTTTATACAACTGGCAGAAGTCATTGTGACATCGAATATCTCCCGCGCGAAATGGA
ACAACACCCTGAGAAACAGATCGTGAAGAAACTTCGAGAACAATTCGGTAATAAACAATCGTATTCAACAACAATCGTATTCAACACTTTTTAATAGCACTTGGAAT
CCCTGAGATAGTTATGCACTCATTCAACTGTGCGGCGAGTTCTTCTGTCGGATTAAGCAGATCATTAATATGTGCAAGAAGTGGGAAAAG
CTATGTACGCCCCGCCTATTCGCGGACAACAATAAGATGCTCTAGTAATATTACCGGATTGTTGCTGACACGCGAGGGAAATAA
TGAAACAGAGATATTTAGACCTGGCGGGAGCCGGGAGCCGGACATGAGAAGGCGACATGAGAAGTAACTGGAGAAGTGAGCTTTACAAATAAAGTCGTAAAGATA
GAACCATTGGGGTAGCAGCCAACCAACCAAAGCAACAACAACTGTTGGGGAATCAACAGCTTCAGGCAAGAGTGCT
ATAATCTTCTCCGAGCGATACCTCAAAGATCAACAACATCTGCTCGGGCATATGGGGATGTTCCGGTAAACTCATATGCACTACCGCCGTGCCC
TGGAACGCGAGCTGGTCTAATAAATCCCTGATGAAATTGGACAACCAACAGGAGAAAAACGAACAGGAACTGCTGGAACTGGACAAGTG
ATACTAGTTTGATTTATACTCTGATCGAAGAATCTCAGAACAACTAACTGGCCTGTGGTAAAGATCTTACAA
```

(For all 140CF design shown here and below, 140CF gene will be flanked with the 5' sequence of "TTCAGTCGACGGCCACC" that contains a Kozak" sequence (GCCACCATGG/A) and SalI site and 3' sequence of TAAAGATCTTACAA containing stop codon and BglII site.)

Fig. 29A

CON OF CON-S-2003 (829 a.a.)
MRVMGIQRNCQHLWRWGILIFGMLIICSAAENLWVTVYYGVPVWKEANTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEIVL
ENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVNATNNTTNNEEIKNCSFNITTEIRDKKKKVYALFYKL
DVVPIDDNNSYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSL
AEEEIIRSENITNNAKTIIVQLNESVEINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNISRTKWNKTLQQVAKKLRE
HFNKTIIFNPSSGGDLEITTHSFNCGGEFFYCNTSELFNSTWNGTNNTITLPCRIKQIINMWQGVGQAMYAPPIEGKIRCTSNIT
GLLLTRDGGNNNTETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKAKRRVVEREKRAVGIGAVFLGFLGAAGSTMGAASITL
TVQARQLLSGIVQQQSNLLRAIEAQQHLLQLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTNVPWNSSWSNKSQDEI
WDNMTWMEWDKEINNYTDIIYSLIEESQNQQEKNEQELLALDKWASLWNWFDITNWLWYIKIFIMIVGGLIGLRIVFAVLSIVNR
VRQGYSPLSFQTLIPNPRGPDRPEGIEEEGGEQDRDRSIRLVNGFLALAWDDLRSLCLFSYHRLRDLILIAARTVELLGRRGWEA

*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF
design and the "W" underlined with red color is the last amino acid at the C
terminus, and all the remaining amino acids after the "W" will be deleted in 140CF
design.

Fig. 29B

CON-S-2003 140CF.pep (620 a.a.).
Nick name: 006
MRVMGIQRNCQHLWRWGILIFGMLIICSAAENLWVTVYYGVPVWKEANTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEIVL
ENVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVNATNNTTNNEEIKNCSFNITTEIRDKKKKVYALFYKL
DVVPIDDNNSYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSL
AEEEIIRSENITNNAKTIIVQLNESVEINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNISRTKWNKTLQQVAKKLRE
HFNKTIIFNPSSGGDLEITTHSFNCGGEFFYCNTSELFNSTWNGTNNTITLPCRIKQIINMWQGVGQAMYAPPIEGKIRCTSNIT
GLLLTRDGGNNNTETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKAKTLTVQARQLLSGIVQQQSNLLRAIEAQQHLLQLTV
WGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTNVPWNSSWSNKSQDEIWDNMTWMEWDKEINNYTDIIYSLIEESQNQQEK
NEQELLALDKWASLWNWFDITNWLW*

*Amino acids seen in blue color is for easy identification of the junction of the
deleted fusion cleavage site.

Fig. 29C

CODON-OPTIMIZED CON-S-2003 140CF.seq (1891 nt

Nick name :006

TTCAGTCGACAGCCACCATGCGGTCCATGGGGATACAGAGGAATTGCCAGCACTTGTGGAGGTGGGGAATTTGATATTCGGGAT
GCTCATAATCTGCTCTGCCCGCTGAGAACCTGTGGGTCACTGTATTACGGCGTTCCCGTCTGAAAGAAGCTAATACTACCCTG
TTTTGTCAAGCGACGCCAAAGCCATACGACACCGAAGTCCACAATGTCTGGGCTACCACGCCTGTGTACCTACTGATCCAAATC
CCCAGGAAATTGTTCTTGAAAACGTAACGGAAAACTTTAACATGTGGAAGAATAATATGGTGGAGCAAAATGCACGAGGATATAAT
CAGCCCTGTGGACCAGTCCCTCAAACCATGCGTTAAACTCACTCACTCTGCGTGACTCTGCGAGTACGGATAAGAAAAAAGGTTTATG
AATAATACAAACAAACAATGAGGAGATAAAGAATTGTTCATTTAATATAACCACTGAGATACGGATAAGAAGAAAAAGGTTTATG
CACTCTTTTACAAGCTCGACGTGGTGCCCATAGACGACAATAATAGCTACCGACCTCATTAATTGCAATACTAGCGCTATAACCCA
GGCATGCCCCCAAAGTTCCTTCGAGCCCATACCGATTCACTACTGCGCACCCGCCGGATTCGCCATTCTTAAATGCAATGACAAG
AAGTTCAACGGCACCGGACCCCTGTAAGAACGTAAGCACTGTTCAATGTACACAGATTCGCAATGTAAGTCGCCGTAGTGTCAACGCAGCTCC
TCCTCAACGGAAGCCTTGCAGAAGAAGAGATCATTATCAGGTCAGAAATAATAACACAAGAAGTCAATTAGGATCGGACCCGGCGAAAAACAATCATTGTTCAGCT
GAATGAGTCTGTAGAAATCAATTGTACCCGCCCTAATAATAACACAAGAAAGTCAATTAGGATCGGACCCGGCGAAAAACAATCATTGTTCTAC
GCAACCGGAGATATCATCGGGGAACATTTAATAAGACAATCATCTTCAATCCAACTGTTCAACTGAGCGGAGGGACCTGGAATAAACTTTGCAGCAGGTAG
CCAAGAAACTGCGGGGCGAGTTTTTCTACTGTAATAAACCTTGTAATCAACATGTGGGGCAAGGTGTGGGGCAAGCAATAACACGGAGACCTTTAGGCCAGGCGGCGGCGATAT
TAACTGTGGGGGCGAGTTTTTCTACTGTAATAACCTTGTAATCAACATGTGGGGCAAGGTGTGGGGCAAGCAATAACACGGAGACCTTTAGGCCAGGCGGCGGCGATAT
CCTTGCAGAATAAAACAGATTATCCGGACTCCTCCTGACTCTGGGGGCAAGGTGTGGGGCAAGCAATAACACGGAGACCTTTAGGCCAGGCGGCGGCTAAAACC
GCACCTCCAATATTACCGGACTCCTCCTGACTCTATAAGTCGTTACGAGCCCGTCTCCTGGAGTTGCGCCAACCAAAGCTAAAACC
GAGAGATAACTGGCGCTCCGAGCTCTATAAATACAAAGTCGTTACAGCAGCAATCTAATCTTTGAGAGCGCTATTGAAAGACCAACAACTTCT
TTGACCGTGCAAGCCAGCAGTTGTTGTCAGGTATCGTACAGCAGCAATCTAATCTTTGAGAGCGCTATTGAAAGACCAACAACTTCT
TCTTGCAGCTTACCGTCTCGGGCCATCAAACAACTTCAGGCACGCGTGTCCTTGGAACAGCAGCAGTCGGTCAAATAAAGCCAAGACGAA
CGGGATCTGGGGGTGTTCTGAAAATTGATCTGCACGACAAATGGCCTTGGAACAGCAGCAGTCGGTCAAATAAAGCCAAGACGAA
ATATGGGGATAACATGACATGGATGGAATGGACAGGAACTCTTGGCTCTGGACAAATGGGCTTCACTGTGGAACTGGTTCGACATCACAAATTG
AAAATCAACAGGAAAAAAAATGAACAGGAACTCTTGGCTCTGGACAAATGGGCTTCACTGTGGAACTGGTTCGACATCACAAATTG
GCTCTGGTAAAGATCTTACAA

Fig. 30A

CONSENSUS A1-2003 (845 a.a.)

MRVMGIQRNCQHLLRWGTMILGMIICSAAENLWVTVYYGVPVWKDAETTLFCASDAKAYETEMHNVWATHACVPTDPNPQEIHL
ENVTEEFNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLNCSNVNVTNNTTHEEEIKNCSFNMTTELRDKKQKVYSLFY
LNGSLAEEEVIIRSENITNNAKTIIVQLTKPVKINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSRSEWNKTLQKVA
KQLRKYFKNKTIIFTNSSGGDLEITTHSFNCGGEFFYCNTSGLFNSTWNNGTMKNTITLPCRIKQIINMWQRAGQAMYAPPIQGV
IRCESNITGLLLTRDGGNNNTNETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRAKRRVVEREKRAVGIGAVFLGFLGAAGS
TMGAASITLTVQARQLLSGIVQQQSNLLRAIEAQQHLLKLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTNVPWNSS
WSNKSQNEIWDNMTWLQWDKEISNYTHIIYNLIEESQNQQEKNEQDLLALDKWANLWNWFDISNWLMYIKIFIMIVGGLIGLRIV
FAVLSVINRVRQGYSPLSFQTHTPNPRGLDRPGRIEEEGGEQGRDRSIRLVSGFLALAWDDLRSLCLFSYHRLRDFILIAARTVE
LLGHSSLKGLRLGWEGLKYLWNLLLYWGRELKISAINLVDTIAIAVAGWTDRVIEIGQRIGRAILHIPRRIRQGLERALL

*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF
design and the "W" underlined with red color is the last amino acid at the C
terminus, and all the remaining amino acids after the "W" will be deleted in 140CF
design.

Fig. 30B

Con-A1-2003 140CF.pep (629 a.a.)
Nick name: 001

MRVMGIQRNCQHLLRWGTMILGMIICSAAENLWVTVYYGVPVWKDAETTLFCASDAKAYETEMHNVWATHACVPTDPNPQEIHL
ENVTEEFNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLNCSNVNVTNNTTHEEEIKNCSFNMTTELRDKKQKVYSLFY
LNGSLAEEEVIIRSENITNNAKTIIVQLTKPVKINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSRSEWNKTLQKVA
KQLRKYFKNKTIIFTNSSGGDLEITTHSFNCGGEFFYCNTSGLFNSTWNNGTMKNTITLPCRIKQIINMWQRAGQAMYAPPIQGV
IRCESNITGLLLTRDGGNNNTNETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRAKTLTVQARQLLSGIVQQQSNLLRAIEA
QQHLLKLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTNVPWNSSWSNKSQNEIWDNMTWLQWDKEISNYTHIIYNLI
EESQNQQEKNEQDLLALDKWANLWNWFDISNWLW*

*Amino acids seen in blue color is for easy identification of the junction of the
deleted fusion cleavage site.

Fig. 30C

CODON-OPTIMIZED Con-A1-2003.seq
Nick name: 001 (1918 nt)

```
TTCAGTCGACAGCCACCATGAGGGTGATGGGAATCCAACGGAACTGCCAGCATCTTCTCCGGTGGGAACGATGATACTGGGAAT
GATAATAATCTGCTCTGCCGCTGAAAACCCTCTGGGTCACAGTGTACTACGGAGTGCCTGTATGGAAGGACGCTGAAACCACTCTC
TTTTGTGCTTCCGATGCTAAAGCCTACAGATGCACAATGTTGGGCCACCCACGCCTGCGTGCCAACTGATCCTAATC
CACAAGAAATACATCTGGAGAATGTTACTGACGAAT|TAACATGTGGAAAAATAATATGGTAGAGCAAATGCACACTGACATCAT
TTCACTCTGGGACCAATCACTCAAACCCTGCGTTAAACTTACCCCCCTCTGCGTGACCCTCAATTGTAGCAACGTCAACGTCACA
AATAATACAACCAACACTCACGAGGAAGAAATTAAAAATTGCTCCTTTAATATGACCACTGAACTTCGGCACAAAAACAAAAAG
TCTATTCACTGTTTTATAGGCTGGACGTCGTCCAAATCAACGAGAACAATTCTAACAGTAGCTATCAATTGCAATAC
CTCTGCTATTACCCAAGCTTGTCCTAAAGTCTCTTTTGAACCAATCCTATCCACTACTGTGCCCCAGCTGGATTCGCAATTCTG
AAGTGCAAGGATAAGGAATTCAACGGAACTGGCCTTCACTCGCCGAAGAAGAGGTGATTATCCGAAGAACATAACAATGCGAAGAC
TCAGCACTCAACTGCTCCTGAATGGCTCAACCAGTGAAGATCAACTGTACTAGACCAAATAACAACAAGAAAATCTATCAGAATTGGCCCC
GGACAAGCCTTCTACGCCAACAGGAGATATCATAGGTGACATCAGAGAACAGCAGGCCCAATCATATTTACTAACGTTTCAAGAAGCGAGTGGAATAAAA
CACTCCAGAAAGTGGCAAGCAGCTTTAATTGTGGGGGCGAATTCTTCTACTGTAACACATCGGCCTTTAATTCTACCTGAATAACGGC
AATAACCACTCATAGCTTTGTGTGGGGGCGAATTCTTCTACTGTAACACATCGGCCTTTAATTCTACCTGAATAACGGC
ACCATGAAAAATACTATCACCCTCCCTTGCAGAATTAAGCAACATTACTGGACTTCTTTCTGACTTGGCGGAAATAATATACCAATGA
CTCCCATTCAAGGTGTGATTCGATGTGAAGAACTGAAGTGCCGAAATGCCGAGACAAATTGGCGAAGTGAACTTTATAAATACAAAGTAGTTAAGATTGAGCCCCTT
GACATTCAGACCCGGCGGCGGCGATATGCCGATATGCCGAAGTGAACTTTATAAATACAAAGTAGTTAAGATTGAGCCCCTT
GGAGTTGCCCCTACTAGACCACAACAACATTGACCGTTCAGGCCAGCAATTGGGGAATCAAGCAATTGCAGGCTAGGGTTTGGCTGTGGA
TCCGAGCTATCGAGGCACCAACAACATCTCTTGGGAATCTGGGGATGCTCTGGGAAAATTGATATGTGCTACAAACGTACCCTGGAACTCA
ACGCTATCTCAAGGATCAGCAGCTTCTCTGGGATCTGGGGATGCTCTGGGAAATTGATATGTGCTACAAACGTACCCTGGAACTCA
AGCTGGGAGTAATAAAGCCAGAACGAAGAATCACAGAACCAGCAGGAAAAGAAATTTCTAATTATACTCATA
TCATATACAATCTGATCGAAGAATCACAGAACCAGCAGGAAAAGAAATTTCTAATTATACTCATA
GTGGAACTGGTTTGACATTAGCAACTGGCTGTGGTGGTAAAGATCTTACAA
```

Fig. 31A

CONSENSUS C-2003 (835 a.a)

```
MRVRGILRNCQQWWIWGILGFWMLMICNVVGNLWVTVYYGVPVWKEAKTTLFCASDAKAYEKEVHNVWATHACVPTDPNPQEIVL
ENVTENFNMWKNDMVDQMHEDIISLWDQSLKPCVKLTPLCVTLNCTNATNATNTMGEIKNCSFNITTELRDKKQKVYALFYRLDI
VPLNENNSYRLINCNTSAITQACPKVSFDPIPIHYCAPAGYAILKCNNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAE
EEIIIRSENLTNNAKTIIVHLNESVEIVCTRPNNNTRKSIRIGPGQTFYATGDIIGDIRQAHCNISEDKWNKTLQKVSKKLKEHF
PNKTIKFEPSSGGDLEITTHSFNCRGEFFYCNTSKLFNSTYNSTNSTITLPCRIKQIINMWQEVGRAMYAPPIAGNITCKSNITG
LLLTRDGGKNNTETFRPGGGDMRDNWRSELYKYKVVEIKPLGIAPTKAKRRVVEREKRAVGIGAVFLGFLGAAGSTMGAASITLT
VQARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQTRVLAIERYLKDQQLLGIWGCSGKLICTTAVPWNSSWSNKSQEDIW
DNMTWMQWDREISNYTDTIYRLLEDSQNQQEKNEKDLLALDSWKNLWNWFDITNWLWYIKIFIMIVGGLIGLRIIFAVLSIVNRV
RQGYSPLSFQTLTPNPRGPDRLGRIEEEGGEQDRDRSIRLVSGFLALAWDDLRSLCLFSYHRLRDFILIAARAVELLGRSSLRGL
QRGWEALKYLGSLVQYWGLELKKSAISLLDTIAIAVAEGTDRIIELIQRICRAIRNIPRRIRQGFEAALQ
```

*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF
design and the "W" underlined with red color is the last amino acid at the C
terminus, and all the remaining amino acids after the "W" will be deleted in 140CF
design..

Fig. 31B

Con-C 2003 140CF.pep (619 a.a.)
Nick name: 003

```
MRVRGILRNCQQWWIWGILGFWMLMICNVVGNLWVTVYYGVPVWKEAKTTLFCASDAKAYEKEVHNVWATHACVPTDPNPQEIVL
ENVTENFNMWKNDMVDQMHEDIISLWDQSLKPCVKLTPLCVTLNCTNATNATNTMGEIKNCSFNITTELRDKKQKVYALFYRLDI
VPLNENNSYRLINCNTSAITQACPKVSFDPIPIHYCAPAGYAILKCNNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAE
EEIIIRSENLTNNAKTIIVHLNESVEIVCTRPNNNTRKSIRIGPGQTFYATGDIIGDIRQAHCNISEDKWNKTLQKVSKKLKEHF
PNKTIKFEPSSGGDLEITTHSFNCRGEFFYCNTSKLFNSTYNSTNSTITLPCRIKQIINMWQEVGRAMYAPPIAGNITCKSNITG
LLLTRDGGKNNTETFRPGGGDMRDNWRSELYKYKVVEIKPLGIAPTKAKTLTVQARQLLSGIVQQQSNLLRAIEAQQHMLQLTVW
GIKQLQTRVLAIERYLKDQQLLGIWGCSGKLICTTAVPWNSSWSNKSQEDIWDNMTWMQWDREISNYTDTIYRLLEDSQNQQEKN
EKDLLALDSWKNLWNWFDITNWLW*
```

*Amino acids seen in blue color is for easy identification of the junction of the
deleted fusion cleavage site.

Fig. 31C

CODON-OPTIMIZED Con-C-2003 140CF (1,888 nt.)
Nick name:003
TTCAGTCGACAGCCACCATGCCAGTGAGTGAGAGGCATTCTGCGGAATTGTCAGCAATGGTGGATCTGGGGCATACTCGGATTCTGGAT
GCTTATGATATGCAATGTTGTGGGAACCTGTGGGTTACCGTATACTATGGGGTTCCAGTCTCGGAAGGAGGCTAAAACAACGCTG
TTCTGTGCAAGTGACGCCAAAGCCTACGAGAAAGAAGTGCACAACGTCTCGGGCTTGTGTTCCAACCGATCCAAACC
CCCAGGAAATCGTCCTCGAGAACGTGACTGAAAACCATGGGAAGAATGATATGGTAGATCAGATGCACGAAGATATCAT
TTCATTGTGGGACCAATCATTGAAAACCATGCGTAAAACTGAAAAAAAAACTGTAGCTTTAACATTACAACGAAACTTAACTGC
ACCAATACTATGGGCGAAATAAAAAAACTGTAGCTTTAACATTACAACGGAACTCCGGGATAAGAAACAAAAGGTCTACGCGCTCT
TTTACCGACTCGATATCGTCCTCCCACTTAACGAGAATAATAGTTACCGCCTGATTAACTGTAACACATCAGCCATTACGCAAGCTTG
CCCAAAGTTCTTTCGACCCCATCCCAATTCACTATTGTGCCCCCGCTGGATACGCTATACTTAAATGCAACAATAAAACATTT
AATGGAACCGGACCATGTAACACGTCAGTACCGTACAATGTACGCACGGAATTAAACCTGTTGTCTCAAACCCAGCTTCTCCTTA
ACGGCTCATTGGCGGAGGAAGAAATTATTATCAGATCAGAAAACTTGACCAACAATGCCAAAACCATCATCGTGCACCTCAATGA
ATCCGTGGAAATCGTGTGCACCAGAGACCAAATAACAATAACCCGGAAATCAATCAGGATTGGGCCTGGCCTGCAGCATTTTACGCTACA
GGTGATATAATTGGCGATATTAGACAAGCCCATTGCAACAATATCGCAACAAGTGGAATAAGACTCTGCAGAAGGTTTCTAAGA
AGCTGAAGGAACACTTTCCCAATAAAAACGATTAAGTTCGAGCCTCTTTAACAGTACCTTCCACTAATAGTACCATCACACTCCCC
TTGTAGAGGGGGAGTTCTTCTATTGTAATACATCAAAGCTCTTTAACAGTACCTTCCACTAATAGTACCATCACACTCCCC
TGCAGAATAAAGCAAATAATCAACATGTGGCAAGAAGTTGGCCGAGCAATAACACTGAGACCTTCAGACCTGGCGAGGCGATATGCG
AATCCAATATTACTGGCCTTTTGCTGACACGGCGAGCGAGCTCTACAAGTATAAAGTCGTTGAAATCAAGCCACTGGGCATAGCTCCTACGAAAGCAAAGACACTC
CGATAATTGGCGCGAGCGAGCGAGCTCTACAAGTATAAAGTCGTTGAAATCAAGCCACTGGGCATAGCTCCTACGAAAGCAAAGACACTC
ACTGTTCAGGTACGAACAGCTGCTCTCCGGCATAGTGCAACAGCAGTCAAGCCCAACATATGC
TCCAGCTTACCGTCTGGGAATCAAACAATTGCAAACAACGAGTGCTGGCGTGAACTCAAGCTATCGAAGCCCAACAACATATGC
GATTTGGGCGTCGTTCAGTGAGATGCGAGTGGGATCGAAGCTCATCGTACAAACAATTGCAAGAGATTTGAAAGATCAGCAACTCCTGGG
TGGGACAACATGACTTGGATGCCAGTGGGATCGAAGCTCATCGTACAACTATACAGATACCCATTTATCGGCTCCTGGAGGACTCACAGA
ACCAGCAGGAGAAAATGAGAGAAAGATTTGTTCGGCGCTTGGAAGAATTTGTGAATTGGTTCGACATTACAAACTGGCT
CTGGTAAAGAGATCTTACAA

Fig. 32A

CONSENSUS G-2003 (842 a.a.)

MRVKGIQRNWQHLWKWGTLILGLVIICSASNNLWVTVYYGVPVWEDADTTLFCASDAKAYSTERHNVWATHACVPTDPNPQEITL
ENVTENFNMWKNNMVEQMHEDIISLWDESLKPCVKLTPLCVTLNCTDVNVTNNNTNNTKKEIKNCSFNITTEIRDKKKKEYALFY
RLDVVPINDNGNSSIYRLINCNVSTIKQACPKVTFDPIPIHYCAPAGFAILKCRDKKFNGTGPCKNVSTVQCTHGIKPVVSTQLL
LNGSLAEEIIIRSENITDNTKVIIVQLNETIEINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSRTKWNEMLQKVK
AQLKKIFNKSITFNSSSGGDLEITTHSFNCRGEFFYCNTSGLFNNSLLNSTNSTITLPCIKQIVRMWQRVGQAMYAPPIAGNIT
CRSNITGLLLTRDGGNNNTETFRPGGGDMRDNWRSELYKYKIVKIKPLGVAPTRARRRVVEREKRAVGLGAVLLGFLGAAGSTMG
AASITLTVQVRQLLSGIVQQQSNLLRAIEAQQHLLQLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTNVPWNTSWSN
KSYNEIWDNMTWIEWEREISNYTQQIYSLIEESQNQQEKNEQDLLALDKWASLWNWFDITKWLWYIKIFIMIVGGLIGLRIVFAV
LSIVNRVRQGYSPLSFQTLTHHQREPDRPERIEEGGGEQDKDRSIRLVSGFLALAWDDLRSLCLFSYHRLRDFILIAARTVELLG
RSSLKGLRLGWEGLKYLWNLLLYWGQELKNSAINLLDTIAIAVANWTDRVIEVAQRACRAILNIPRRIRQGLERALL

*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF
design and the "W" underlined with red color is the last amino acid at the C
terminus, and all the remaining amino acids after the "W" will be deleted in 140CF
design.

Fig. 32B

Con-G-2003 140CF (626 a.a.)
Nick name: 007

MRVKGIQRNWQHLWKWGTLILGLVIICSASNNLWVTVYYGVPVWEDADTTLFCASDAKAYSTERHNVWATHACVPTDPNPQEITL
ENVTENFNMWKNNMVEQMHEDIISLWDESLKPCVKLTPLCVTLNCTDVNVTNNNTNNTKKEIKNCSFNITTEIRDKKKKEYALFY
RLDVVPINDNGNSSIYRLINCNVSTIKQACPKVTFDPIPIHYCAPAGFAILKCRDKKFNGTGPCKNVSTVQCTHGIKPVVSTQLL
LNGSLAEEIIIRSENITDNTKVIIVQLNETIEINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSRTKWNEMLQKVK
AQLKKIFNKSITFNSSSGGDLEITTHSFNCRGEFFYCNTSGLFNNSLLNSTNSTITLPCIKQIVRMWQRVGQAMYAPPIAGNIT
CRSNITGLLLTRDGGNNNTETFRPGGGDMRDNWRSELYKYKIVKIKPLGVAPTRART**LLQLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTNVPWNTSWSNKSYNEIWDNMTWIEWEREISNYTQQIYSLIEES
QNQQEKNEQDLLALDKWASLWNWFDITKWLW***

*Amino acids seen in blue color is for easy identification of the junction of the
deleted fusion cleavage site

Fig. 32C

CODON-OPTIMIZED Con-G-2003 140CF.seq
Nick name:007

TTCAGTCGACAGCAGCCACCACCATGCCAGCTGAGTCGAGGGAATCCAGAGAAATTGGCAGCAGCACCTTTGGAAGTGGGCCACACTCATCCTCGGCCT
TGTGAICATATGCTCTGCCTCAAATAACCTTTGGGTCACAGTTTATTACGGCGTGCCCGTTTGGGAGGACGGCAGACACAACTCTT
TTTTGCCAGCGACGCTAAGGCTTATTCAACAGAGAGGCATAACGTTTGGGCTACACATGCCATGCCGTGCCGACCGATCCTAATC
CCCAGGAAATCACTCTTGAGAATGTTACAGAGAATTTAATATGTGGAAGAACAACATGGTTGAACAGATGCATGAAGACATAAT
TTCTCTCTGGGATGAATCTCTGAAACCTTGCGTGAAGCTTACCACCACTGTGCCTGAATTGCACTGACGTCAATGTCACA
AATAATAATACCAACACAATACAAAAAAGAAATCAAAAAATTGTTCTTTCAACATAACCAGAGATACGCGATAAAAAAGAAAG
AATACGCCCCTGTTCTACAGACTCGATGTGGTCCCAATTAATGACAACGGAAATTCTTCCATCTACCGACTTATCAATTGTAACGT
GTCTACAATCAAAACAGGCCTGTCCTAAAGTCACATTGACCCTATTCCCATTCATTACTGTGCCCCGCTTCGCTATTCTT
AAATGCCGAGACAAAAAATTAACGGACAGACCATGCAAGAAGTGTCTCAACAGTTCAATGCACTCATGGAATTAAACCAGTCG
TTTCTACTCAACTCCTTCTCAATGGAGCGCTGAAATAAATTGTACGAGACCTAATAACAACAAGAAAACATAACCGACAACACAAAAGT
AATCATCGTACAGCTGAACGAGACCATTGAAATAAATTGTACGAGACATATCCGCCAGCTCAACAAATCTATTCAACAGAAAATTATT
GGACAGGCTTTCTACGCCACAGGAGACATTATCCGGAGATATCCGCCAGCTCAACAAATCTATTCAACAGAAAATTATTCAACAGAAAATTATTCAACAATCTATT
TGCTTCAGAAGGTCAAAGCTCAGCTCAAGAAGCTCAAGAAATATTCAACAAATCTATTCAACAATCTATTCAACAGAAAATTATT
AACAACTCATTCCTTCAACTGTCGGGGAGAATTTTTTACTGTAACACGTCCGGCTCGTTCAACAATTCACTCCTGAATAGCACT
AACTCCACCATCACTCTCCCATGTAAGATCAAAACAAATCGTCAGAATGTGGCAGCGAGTCGGCTCAGCTATGTACGCCCCTCCAA
TCGCCGGTAATATCACATGTAGAAGCAATATCACAGGGCTCTTGCTTACAAGGGGAGCCGAGCTGTATAAATATAAGATCGTAAAAATATAAGACG
ACCAGGAGGAGGAGAGACATGCCGAGACAATTGGCGAGGTCAGCAACTGCTAGGGAATCAAACAACTCAATCTTCTTAGAGCAA
CCAACTAGAGCCCGAACACTGACCGTGCAGGTGAGGCAACTGCTAGGGAATCAAACAACTGCTATTGGCAGTGGTGTAGCG
TCGAGGCCCAGCAGCATCTGCTCCAGCTTACTGTATGGGAATCAAACAACTGCTATTGGCAGTGGAGGTATCT
CAAGGACCAGCAGCTTCTGCTCCAGCTTACTGTATGGGGACAATATGGGGACAATATGCAGTGGAATATCCAGCAAATCTATT
AATAAGAGTTACAATGAAATCTGGGACAATATGGGGACAATATGACATGGGAGCGCGAAATATCCAACTATACTCAGCAAATCTATT
CCCTCATTGAAGAGAGTCAGAACCAGCAGCAAAAGAATGAGCAAGAACCCTCCTGCGCCCTGGATAAATGGGCATCTCTGTGAACTG
GTTTGACATAACTAAATGGTTGTGTAGAGATCTTACAA

SUBSTITUTE SHEET (RULE 26)

Fig. 33A

CONSENSUS_01_AE-2003 (854 a.a.)
MRVKETQMNWPNLWKWGTLILGLVIICSASDNLWVTVYYGVPVWRDADTTLFCASDAKAHETEVHNVWATHACVPTDPNPQEIHL
ENVTENFNMWKNNMVEQMQEDVISLWDQSLKPCVKLTPLCVTLNCTNANLTNVNNITNVSNIIGNITNEVRNCSFNMTTELRDKK
QKVHALFYKLDIVQIEDNNSYRLINCNTSVIKQACPKISFDPIPIHYCTPAGYAILKCNDKNFNGTGPCKNVSSVQCTHGIKPVV
STQLLLNGSLAEEEIIIRSENLTNNAKTIIVHLNKSVEINCTRPSNNTRTSITIGPGQVFYRTGDIIGDIRKAYCEINGTKWNEV
LKQVTEKLKEHFNNKTIIFQPPSGGDLEITMHHFNCRGEFFYCNTTKLFNNTCIGNETMEGCNGTIILPCKIKQIINMWQGAGQA
MYAPPISGRINCVSNITGILLTRDGGANNTNETFRPGGGNIKDNWRSELYKYKVVQIEPLGIAPTRAKRRVVEREKRAVGIGAMI
FGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLRAIEAQQHLLQLTVWGIKQLQARVLAVERYLKDQKFLGLWGCSGKIIC
TTAVPWNSTWSNRSFEEIWNNMTWIEWEREISNYTNQIYEILTESQNQQDRNEKDLLELDKWASLWNWFDITNWLMYIKIFIMIV
GGLIGLRIIFAVLSIVNRVRQGYSPLSFQTPTHHQREPDRPERIEEGGEQGRDRSVRLVSGFLALAWDDLRSLCLFSYHRLRDF
ILIAARTVELLGHSSLKGLRRGWEGLKYLGNLLLYWGQELKISAISLLDATAIAVAGWTDRVIEVAQGAWRAILHIPRRIRQGLE
RALL

*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF
design and the "W" underlined with red color is the last amino acid at the C
terminus, and all the remaining amino acids after the "W" will be deleted as 140CF.

Fig. 33B

Con-AE01-2003 140CF.pep (638 a.a.)
Nick name: 008
MRVKETQMNWPNLWKWGTLILGLVIICSASDNLWVTVYYGVPVWRDADTTLFCASDAKAHETEVHNVWATHACVPTDPNPQEIHL
ENVTENFNMWKNNMVEQMQEDVISLWDQSLKPCVKLTPLCVTLNCTNANLTNVNNITNVSNIIGNITNEVRNCSFNMTTELRDKK
QKVHALFYKLDIVQIEDNNSYRLINCNTSVIKQACPKISFDPIPIHYCTPAGYAILKCNDKNFNGTGPCKNVSSVQCTHGIKPVV
STQLLLNGSLAEEEIIIRSENLTNNAKTIIVHLNKSVEINCTRPSNNTRTSITIGPGQVFYRTGDIIGDIRKAYCEINGTKWNEV
LKQVTEKLKEHFNNKTIIFQPPSGGDLEITMHHFNCRGEFFYCNTTKLFNNTCIGNETMEGCNGTIILPCKIKQIINMWQGAGQA
MYAPPISGRINCVSNITGILLTRDGGANNTNETFRPGGGNIKDNWRSELYKYKVVQIEPLGIAPTRAKTLTVQARQLLSGIVQQQ
SNLLRAIEAQQHLLQLTVWGIKQLQARVLAVERYLKDQKFLGLWGCSGKIICTTAVPWNSTWSNRSFEEIWNNMTWIEWEREISN
YTNQIYEILTESQNQQDRNEKDLLELDKWASLWNWFDITNWLW*

*Amino acids seen in blue color is for easy identification of the junction of the
deleted fusion cleavage site.

SUBSTITUTE SHEET (RULE 26)

Fig. 33C

CODON-OPTIMIZED Con-AE01-2003 140CF.seq (1945 nt.)

Nick name: 008

```
ttcagtcgacagccaccATGCGCAGCCAGAGTCAAGGAAACACAAATGAACTGGCCTAATCTGTGGAAGTGGGCACCCTGATCCTGGGTTT
GGTCATTATTTGCTCTGCGAGCGACAATCTCTGGGTTACTGTCTATTACGGAGTCCCCGTTTGGAGAGATGCCGACACTACACTG
TTCTGCGCCTCAGATGCCAAAGCTCATGAAACTGAAGTGCATAATGTTTGGGCAACCCACGCCTGTGTTCCTACCGACCCAAACC
CCCAAGAAATACACCTGGAAAAACGTGACCGAGAACTTTAATATGTGGAAGAATAACATGGTTGAACAGATGCAAGAAGACGTAAT
CAGCCTGTGGGATCAAAGTCTGAAAACCTTGCGTAAAACTGACTCCACTTTGCGTAACACTTAATTGCACCACCGAACCTGACA
AACGTTAACAACATCACTAACGTGTCTCCAACATCATCGGCAACATAACGAAGCGAAGTGAGAAATTGCAGTTTCAATATGACTACAG
AGCTCCGGGACAAGAAAACAGAAGGTCCATGCTGCTCTTTTACAAACTCGACATCGTCCAGATCGAAGACAATAACAGCTACAGACT
TATAAATTGTAATACATCCGTGATTAAACAGCATGCCCCCAAAATAAGCTTCGATCTTCCTATCCACTACTGTACTCCTCGCC
GGCTATGCTATCTTGAAATGCAATGGACACCCAGCTGCTCCTCGAACGGCTCTCTGCGAGAGGAAGAGATTATTCGAAGTGAGAACCTCAC
CAACAACGCTAAGACTATCATCGTACATCTCAATAAATCAGTCGAAATTAATTGCACCAGACCCTCCAATAATACTAGAACTTCA
ATCACTATCGGCCCAGGACAAGTCTTTATAGAACAGGAGATATCATAGGGAAGCTTAAGGAACATTCAACAATAAAAACCATTATTTTCAACCCCCAG
CAAAATGGAACGAAGAAGTACTCACCACTGCCACCATGCCAGAAATCACTATGAAGGGCTGCAATGGAACAATCATATCTCAAGATAAAAAATCATTAACATGTGGC
ACGTGCATCGGAAATGAGAGACCATGGAGGGCTGCAATGGAACAATCATATCTCAAGATAAAAAATCATTAACATGTGGC
AAGGTGCTGGACAAGAGCTATGTATGCGACCAGGGCGGGCAACATTAAGGACAACTGGCGGTCCGAACTCTATAAG
AGACGGGAGGAGCAAAATAATACAAATGAAACATTCCGACCAGGACTCGCCGCTAAGAACAACATCTGCTTGGTCTCTGGGATGTTCTGGCAAATAATC
TACAAAGTCGTACAGCAGCAATCCAACCTCTTGGAATAGCACCGAGGCCCAACAACACATCTGCTTGGTCTCTGGGATGTTCTGGCAAATAATC
GAATCGTACAGCAGCAATCCAACCTCTTGGAATAGCACCGAGGCCCAACAACACATCTGCTTGGTCTCTGGGATGTTCTGGCAAATAATC
GCTCCAGGCACGCGTGCTCGCAGTGGAAAGATACCTGAAGGATCAGAAATTCCTTGGTCTCTGGGATGTTCTGGCAAAATAATC
TGCACTACCGCGTTCCCTGGAATTCAACATGACGAACCAGATATACGAAATCCTCACCGAAAGCCAAAATCCAGATGGAACAATAATC
AAAGGGAAATTAGTAACTATACGAAACCAGATATACGAAATCCTCACCGAAAGCCAAAATCCAGATGATAGAGTGGG
CCTCGAGCTTGATAAGTGGCATCCCTTTGGAACTGTTCGACATCGGTTCGACATCATCACAAATTGGCTCTGGTaaagatcttacaa
```

Fig. 34A

Wild-type subtype A Env

00KE_MSA4076-A (Subtype A, 891 a.a),

MGAMGIQMNWQNLWRWGTMILGMLIICSVAEKSWVTVYYGVPVWRDAETTLFCASDAKAHDKEVHNVWATHACVPTDPNPQEMIL
ENVTEDFNMWKNSMVEQMHTDIISLWDQSLKPCVKLTPLCVTLNCSDSNITSNSTKDSATLDMKSEIQNCSFNMTTELRDK
KQKVYSLFYRLDVVQINENSSDYRLINCNTSAITQACPKVTFEPIPIHYCAPAGFAILKCNDKKFNGTGPCTNVSTVQCTHGIKP
VVTTQLLLNGSLAEEEVMIRSENITENAKNIIVQFKEPVQIICIRPGNNTRKSVHIGPGQAFYATGDIIGDIRQAHCNVSRELWN
KTLQEVATQLRKHFRNNTKIIFTNSSGGDVEITTHSFNCGGEFFYCDTSGLFNSSWTASNDSMQEAHSTESNITLQCRIKQIINM
WQRAGQAMYAPPIPGIIRCESNITGLILTRDGGEGNNSTNETFRPVGGNMRDNWRSELYKYKVKVEPLGVAPTKSRRRVVEREK
RAVGLGAVFIGFLGAAGSTMGAASMTLTVQARQLLSGIVQQQSNLLRAIEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGI
<u>WGCSGKLICTTNVPWNSSWSNKSLDEIWENMTWMQWDKEVSNYTQMIYNLLEESQNQQEKNEQELLALDKWANLWN</u>WENISNWLW
YIKIFIMIVGGLIGLRIVFAVLSVINRVRQGYSPLSFQTHTPNPRGLDRPGRIEEGGEQDRDRSIRLVSGFLALAWDDLRSLCL
FSYHRLRDFILIAARTLELLGHNSLKGLRLGWEGLKYLWNLLAYWGRELKISAISLVDSIAIAVAGWTDRIIEIVQAIGRAILHI
PRRIRQGLERALI

*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF
design and the "W" underlined with red color is the last amino acid at the C
terminus, and all the remaining amino acids after the "W" will be deleted in 140CF
design.

Fig. 34B

00KE_MSA4076-A 140CF.pep (647 a.a)

Nick name: 011

MGAMGIQMNWQNLWRWGTMILGMLIICSVAEKSWVTVYYGVPVWRDAETTLFCASDAKAHDKEVHNVWATHACVPTDPNPQEMIL
ENVTEDFNMWKNSMVEQMHTDIISLWDQSLKPCVKLTPLCVTLNCSDSNITSNSTKDSATLDMKSEIQNCSFNMTTELRDK
KQKVYSLFYRLDVVQINENSSDYRLINCNTSAITQACPKVTFEPIPIHYCAPAGFAILKCNDKKFNGTGPCTNVSTVQCTHGIKP
VVTTQLLLNGSLAEEEVMIRSENITENAKNIIVQFKEPVQIICIRPGNNTRKSVHIGPGQAFYATGDIIGDIRQAHCNVSRELWN
KTLQEVATQLRKHFRNNTKIIFTNSSGGDVEITTHSFNCGGEFFYCDTSGLFNSSWTASNDSMQEAHSTESNITLQCRIKQIINM
WQRAGQAMYAPPIPGIIRCESNITGLILTRDGGEGNNSTNETFRPVGGNMRDNWRSELYKYKVKVEPLGVAPTKSR**TLTVQARQ**
LLSGIVQQQSNLLRAIEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICTTNVPWNSSWSNKSLDEIWENMTW
MQWDKEVSNYTQMIYNLLEESQNQQEKNEQELLALDKWANLWNWENISNWLW*

*Amino acids seen in blue color is for easy identification of the junction of the
deleted fusion cleavage site.

Fig. 34C

CODON-OPTIMIZED OOKE_MSA4076-A 140CF.seq (1972 nt.)

Nick name: 011

```
ttcagtcgacagccaccATGGGGGCAATGGGAATCCAGATGAACTGGCAGAACCTCTGGCGATGGGGCACAATGATCCTGGGTAT
GCTCATCATCTGCTCTGTTGCAGAAAAGTCATGGGTAACAGTCTACTACGGCGTACCAGTGTGGCGGGACGCCGAAACCACTCTC
TTCTGCCGCCTCCGATGCCAAAGCCACACGATAAAGAAGTCCACAATGTTTGGGCTACCCATGCGCGTGCCCAACCGATCCTAACC
CACAAGAAATGATACTCGAAAACGTTACTGAAGACTTCAACATGTTGAAAAAATTCTATGGTTGAACAGATGCACACCGACATAAT
ATCACTGTGGGATCAGTCTGTCAAACCCTGTGTCAAATTGACCCCCCTCTGCGTTACACTGAACTGTTCCGACTCAAATATCACT
TCTAATTCAACGAGCAATAGTACGAAAAGACTCCGCAACCTGATATGAAAAGCGAAATACAGAACTGTTCATTTAATATGACCA
CCGAACTGAGAGAGATAAAAAGCAGAAGGTTTATTCTCTGTTCTATCGATTGACGTGGTTCAGATTAACGAAAATAGCAGCGATTA
CCGACTCATTAACTGCAATACATCAGCAATCACACAGGCTTGCCCAAAGTAACATTTGAGCCAATCCCTATTCACTACTGCGCC
CCTGCAGGATTTGCCATCCTGAAATGCAACGATAAGAAGTTTAATGGGACAGCACCCTGCTGAAGAGGAAGTCATGATTCGGTTCTGAAAAA
CCCACGGCATAAAACCTGTTGTTACCACACAAATATTATAGTTCAGTTCAAAGAACCCGTCCAGATCATTGCATTGCCCTGGTAACACTCGC
AAGTCAGTGCACATTGGGACCCAGGCTTTGCAGGAAGTCACTACCCGTCATTAACTGTGGCGGAGAATTCTTCATTGCGAGATACCTCTGGGCTC
GCCGGGAATTGTGGAACAAACTTTGCAGGATCACTACCCATTCAATGCAACAGTAGTATATCAGTGCCAGTAATACAGTA
TAATTCATCAGGCGGTGACGTGGGAGATCACTGCTAGCAACGATTCAATGCAACAATGAAGTACGCACCTCTACAAACGAGACTTTCTACAACGAGACTTTCTACAACGAGAATAATATCGAGAGGCAATATGCGA
TTTAATTCCTCAGCAAGTTGCAGGAAGTTCAAGAAGCACATTCCACAGAAAGTAATATCACACTGCAGTGCCGAATTA
AACAAATCATCAATATGTGGCACCCGAGACGTGGCGAAGGTAATAATTCTACAAACGAGACTTTCAGACCCGTAGGAGGCAATATGCGA
CACTGGCCTCATTCTGACCCGAGACGTGGCGAAGGTAATATAAGTGGTGAAGCTGCCAACAGCAATCCAATCTTCTGAACGAAGTTCCAACAGCAATCCAATCTTCTGAAGGTAGAACCTTCTTGGAGTGGCACCCACCAAATCACGAACCCTGA
CTGTGCAGGCACGCCAACTTCTGAGCGGAATAGTCCAACAGCAATCCAATCTTCTGAACGAAGTTCCAACAGCAATCCAGCAACACCTGCT
TAAACTTACGGTGTGGGGGATGTTCCGGTAAGTTGATTTGACGACAAAACGTTCCCTGGAACTCTTCCTGGTCAAACAGAGTCTGACGAAATAT
GGGAAAATATGACATGGATGGCAGTGGACAAGGAAGTTAGCAACTATACACAGATGATCTACAACCTCCTGAAGAATCTCAGAA
TCAACAGGAAAAAACGAACAAGAACTGCTCGCCCTCGATAAGTGGGCTAACCTCTGGAACTGTTTAATATTTCAAACTGGTTG
TGGtaaagatcttacaa
```

Fig. 35A

Wild-type subtype B
QH0515.1g gp160 (861a.a)

MRVKEIRRNCQRLRRWGTMLLGMLMICSATEQLWVTVYYGVPVWKEATTTLFCASDAKAYVTEKHNVWATHACVPTDPNPQEVVL
ENVTENFNMWKNNMVEQMHEDIISLWEQSLKPCVKLTPLCVTLNCTDKLRNDTSGTNSSSWEKVQKEIKNCSFNITTGIRGRVQ
EYSLFYKLDVIPIDSRNNSNNSTEFSSYRLISCNTSVITQACPKISFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQCT
HGIKPVVSTQLLLNGSLAEEEVVIRSENFTNNVKSIIVQLNKSVVINCTRPNNNTRKSIHIGAGKALYTGEIIGDIRQAHCNLSR
IVNMWQKVGKAMYAPPIRGQIRCSSKITGLILTRDGGTNGTNETETFRPGGGNMKDNWRSELYKYKVVKIEPLGIAPTKAKRRVV
QREKRAVGTIGAMFLGFLGAAGSTMGAASLTLTVQARLLLSGIVQQQNNLLRAIEAQQHLLQLTVWGIKQLQARVLAVERYLRDQ
QLLGIWGCSGRLICTTNVPWNTSWSNRSLNYIWDNMTWMQWDREINNYTDYIYTLLEDAQNQQEKNEQELLELDKWASLWNWFDI
TNWLWYIKIFIMIVGGLIGLRIVFAVLSIVNRVRQGYSPLSLQTHLPARRGPDRPEGIEEGGERDRDRSVRLVHGFLALVWEDL
RSLCLFSYHRLRDLLLIVARTVEILGQRGWEALKYWWNLLLYWSLELKNSAVSLVDTIAIAVAEGTDRIIEIARRIFRAFLHIPT
RIRQGLERALL

*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF
design and the "W" underlined with red color is the last amino acid at the C
terminus, and all the remaining amino acids after the "W" will be deleted in 140CF
design

Fig. 35B

QH0515.1g 140CF (651a.a)
Nick name: 012

MRVKEIRRNCQRLRRWGTMLLGMLMICSATEQLWVTVYYGVPVWKEATTTLFCASDAKAYVTEKHNVWATHACVPTDPNPQEVVL
ENVTENFNMWKNNMVEQMHEDIISLWEQSLKPCVKLTPLCVTLNCTDKLRNDTSGTNSSSWEKVQKEIKNCSFNITTGIRGRVQ
EYSLFYKLDVIPIDSRNNSNNSTEFSSYRLISCNTSVITQACPKISFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQCT
HGIKPVVSTQLLLNGSLAEEEVVIRSENFTNNVKSIIVQLNKSVVINCTRPNNNTRKSIHIGAGKALYTGEIIGDIRQAHCNLSR
AQWNNTLKQIVIKLREQFGNKTIVFNQSSGGDVEIVMHSFNCGGEFFYCNSTQLFNSTWNGNDTWNDTTNDNITLPCRIKQ
IVNMWQKVGKAMYAPPIRGQIRCSSKITGLILTRDGGTNGTNETETFRPGGGNMKDNWRSELYKYKVVKIEPLGIAPTKAKTLTV
QARLLLSGIVQQQNNLLRAIEAQQHLLQLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGRLICTTNVPWNTSWSNRSLNYIWD
NMTWMQWDREINNYTDYIYTLLEDAQNQQEKNEQELLELDKWASLWNWFDITNWLW*

*Amino acids seen in blue color is for easy identification of the junction of the
deleted fusion cleavage site.

Fig. 35C

CODON-OPTIMIZED QH0515.1g 140CF.seq (1984 nt.)
Nick name:012

```
ttcagtcgacagccaccATGAGAGTAAAAGAAATCAGACGCCAACTGTCAGAGGTTGAGGAGGATGGGAACGATGCTCCTGGCAT
GCTGATGATTTGCAGTGCCACCGACAGCTTTGGGTAACCGTGTACCTGTGTATGGTGTACCTGTATGGTGAAAGAAGCCACTACAACCCTG
TTTTGCGCGTCCGACGCAAAAGCCTACGTAACAGAAAAGCCTACGTAACAGAAAATTTAATATGTGAAAAACAATATGGTGAAGATATCAT
CTCAGGAAGTCGTTCTGGAAAATGTAACAGAAAATTTAATATGTGAAAAACAATATGGTGAGACGAGATGCATGAAGATATCAT
CTCACTGTGTGGGAACAATCCTTGAAACCTTGTGTCAAACTGACCCCCACTTTGCGTAACACTTAACTGTACTGATAAGCTTCGCAAT
GATACGTCCGAACAAATTCAAGCAGCTGGGAAAAGGGCGAAATCAAAAATTGTTCATTTAACATCACTACCGGTA
TCAGAGGGCGGGTACAGGAATATTCTCTTTTCTACAAACTCGACGTCATCCCAATCCCAGAAATAACTCAAATAATAGCAC
AGAATTTAGTAGTTATCGCCTTATAAGCTGCAACACCAGCGTGATTACAACGACAAGAAATTTAACGAACCGACCCTGTAAGAATGTGT
ATTCACTACTGCGCACCAGCCGGCTTCGCCATCCTCAAATGTAACGACAAGAAATTTAACGAACCGACCCTGTAAGAATGTGT
CCACCGTTCAATGCACTCATGGAATCAAGCCCGTCGTTCTCCAGCTTAATAAATCCGTCGTTATTAATTGTACAAGACCC
GATTCGCTCCGAAAATTTTACAAACACCATTCACATAGGGCCGGGAAAGCTCTGTATACCGGGAAATTATTGGAGACATCAGACAAGCAC
AACAATAAACACCAGAAAAATCCATTCACATAGGGCCGGGAAAGCTCTGTATACCGGGAAATTATTGGAGACATCAGACAAGCAC
ACTGTAACTTGAGTCGCGCCCAGTGGAACAACACATTGAAAACAGATCGTAATGACACCTGGAAAGATACGACACCACCTATAAGACTAT
CGTGTTTAATCAGAGCTCCGGCGGTGATGTCGAAATCGTAATGACACCTGGAAAGATACGACACCACCTATAAGACTAT
ACACAATTGTTAACAGCACCTGTAAATCGTAAATATGTGGCAAGAAAGTGGCAAGGCCACGGGAACGGGACCTTCCGACCAGGAGGC
CGTGCAGAATAAAGCAAATCGTAAATATGTGGCAAGAAAGTGGCAAGGCCACGGGAACGGGACCTTCCGACCAGGAGGC
TTCTTCCAAGATCACAGGTCTGATACTCACAGTGGCAAGAAAGTGGCAAGGCCACGGGAACGGGACCTTCCGACCAGGAGGC
GGCAACATGAAGGATAACTGGAGAAGTGAACTTTACAAGTATAAAGTGGTCAAGAACAACCTTCTTAGAGCCATTGAAGCACA
ACAACACTGCTGCTCAGTTGACAGTGTGGGGAATTAAAACAGTTGCAGGCCGCCGCCTAAAAATGTCCCTTGGAACGGTATCTTAGAGATCAG
CAGCTTTTGGGTATCTGCAGGTTGACAGTGTGGGGAATTAAAACAGTTGCAGGCCGCCGCCTGGAACGGTATCTTAGAGATCAG
TTAATTATATTGGGACAATATGACATGACATGGGATAGAGAAATTAATAACTACACCGACTACATCTACACACTTCTGGA
GGACGCCCAGAATCAGCAGGAGGAGAACGAGCAGGAACTCCCTCGAATTGGATAAGTGGGCATCACTGTTGGTTCGATATA
ACTAATTGGCTTTGGtaaagatcttacaa
```

Fig. 36A

Wild-type subtype C
DU123.6 gp160(854 a.a)
MRVKGIQRNWPQWWIWGILGFWMIIICRVVGNLWVTVYYGVPVWTEAKTTLFCASDAKAYEREVHNVWATHACVPTDPNPQEIVL
GNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLNCTDVKVNATSNGTTYNNSIDSMNGEIKNCSFNITTEIRDK
KQKVYALFYRPDVVPLNENSSSYILENCNTSTTQACPKVSFDPIPIHYCAPAGYAILKCNNKTFNGTGPCHNVSTVQCTHGIKP
VVSTQLLLNGSLAEEEIIIRSENLTNNAKTIIVHLNESIEIVCTRPNNNTRKSIRIGPGQTVYATNDIIGDIRQAHCNISKTKWN
TTLEKVKEKLKEHFPSKAITFQPHSGGDLEVTTHSFNCRGEFFYCDTTKLFNESNLNTTNTTLTLPCRIKQIVNMWQGVGRAMY
APPVEGNITCNSSITGLLLVRDGGNTSNSTPEIFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTKAKRRVVEREKRAVGIGAVL
FGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLIC
PTTVPWNSSWSNKSQTDIWDNMTWMQWDREISNYTGTIYKLLEESQNQQEKNEKDLLALDSWKNLWSWFDITNWLWYIKIFIMIV
GGLIGLRIIFGVLSIVKRVRQGYSPLSFQTLTPNPRGLDRLGRIEEGGGEQDKDRSIRLVNGFLALAWDDLRSLCLFSYHRLRDF
ILVAARAVELLGRSSLRGLQRGWEALKYLGNLVQYGGGLELKRRAISLFDTIAIAVAEGTDRILEVILRIIRAIRNIPTRIRQGFE
AALL

Fig. 36B

DU123.6 140CF (638 a.a)
Nick name: 013
MRVKGIQRNWPQWWIWGILGFWMIIICRVVGNLWVTVYYGVPVWTEAKTTLFCASDAKAYEREVHNVWATHACVPTDPNPQEIVL
GNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLNCTDVKVNATSNGTTYNNSIDSMNGEIKNCSFNITTEIRDK
KQKVYALFYRPDVVPLNENSSSYILINCNTSTTQACPKVSFDPIPIHYCAPAGYAILKCNNKTFNGTGPCHNVSTVQCTHGIKP
VVSTQLLLNGSLAEEEIIIRSENLTNNAKTIIVHLNESIEIVCTRPNNNTRKSIRIGPGQTVYATNDIIGDIRQAHCNISKTKWN
TTLEKVKEKLKEHFPSKAITFQPHSGGDLEVTTHSFNCRGEFFYCDTTKLFNESNLNTTNTTLTLPCRIKQIVNMWQGVGRAMY
APPVEGNITCNSSITGLLLVRDGGNTSNSTPEIFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTKAKTLTVQARQLLSGIVQQQ
SNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTTVPWNSSWSNKSQTDIWDNMTWMQWDREISN
YTGTIYKLLEESQNQQEKNEKDLLALDSWKNLWSWFDITNWLW*

*Amino acids seen in blue color is for easy identification of the junction of the deleted fusion cleavage site.

Fig. 36C

CODON-OPTIMIZED DU123.6 140CF.seq (1945 nt.)

Nick name: 013

```
ttcagtcgacagccaccATGCGCGTAAAGGGGATTCAAAGAAATTGGCCGCAATGGTGATTTGGGAATTCTGGGCTTTTGGAT
GATAATTATATGCCGCGTTGTCGGAAATTTGTGGGTGACTGTGTACTACGGGGTGCCCGTGTGGACTGAGGCAAAGACCACCCTG
TTCTGTGCTAGCGATGCCAAAGCCTATGAACGCGAAGTGCACAATGTTTGGGCTACTCATGCCTGTGTCCCTACCGACCCAAACC
CTCAGGAAATAGTGCTCGGCAATGTAACGGAAAACTTCAACATGTGGAAAAATGATATGGTGGATCAGATGCACGAAGACATTAT
CTCAATCTGGGACCAAAGCCTGAAACCCTGCGTTAAACTGACTCCTCTCTGCGTTAAACTCAATTGCACAGATGTCAAAGTGAAT
GCCACCTCAAACGGTACGACAACTTACACACAATTCTATTGACTCTATGAACGGCGAAATCAAAAATTGTTCCTTTAACATCACCA
CCGAGATACGCGACAAAAGCAGAAGGTCTATGCCCTTTTTACGCCCGACGTAGTCCCACTCAACGAGAATTCCAGCTCATA
CATCCTCATCAACTGCAATACATCAACTCACCACACAAGCATGCCCGAAAGTTAGCTTTGATCCAATTCCTATACATTACTGCGCC
CCCGCCGGCTACGCTATACTGAAATGCAATAATAAGACTTTTAACGGACCGGCCCATGTCAGAAGAGAAATTATTATCCGCTCGAGAA
CTCATGGCATCAAGCCCGTGGTGTCAACCCAGCTGCTCACCTTAATGAATCAATAGAAATCGTGTGTACTCGGCCCAACAATAATACTAGA
TCTTACTACAATGCAAAAACGATTATCGTGCACCTTACGCAACAGTTTACGCAACTAATCATCATCGGGACACTCCGACAGGCCCATTGCAACATTT
AAAAGCATTCGCATCGGACCTGGCCAGACAGTTTACGCAACAGTTTACGCAACTAATGACATCATCGGGACACTCCGACAGGCCCATTGCAACATTT
CTAAAACCAAGTGGAATACAACCCTGGAAGTCACACCACAAATACAACCACACTGACCCTCCCCGTCGAAGGAAATATAACGTGTAACAGCAGCATCACTGGGCGCTTCTTGTTCGAGACGG
TCACAGTGGCGGGAGACTTGGAGTCACACCACAAATACAACCACACTGACCCTCCCCGTCGAAGGAAATATAACGTGTAACAGCAGCATCACTGGGCGCTTCTTGTTCGAGACGG
AATGAATCAAATCTCAACACCACAAATACAACCACACTGACCCTCCCCGTCGAAGGAAATATAACGTGTAACAGCAGCATCACTGGGCGCTTCTTGTTCGAGACGG
TTGGAAGGGCTATGTACGCTCCCCCCGTCGAAGGAAATATAACGTGTAACAGCAGCATCACTGGGCGCTTCTTGTTCGAGACGG
AGGCAATACTTCTAATTCAACTCCTGAAATTTTAGGCCTGGCGGTGGCAATATGAAAGATAACTGGCGCTCAAGCAAGACAGCTCCTTTCAG
TACAAAGTTGTTGAAATTAAGCCCCTGGGAGTCGCTCCAACCAAAGCCCCAACAGCCCAACACGCATATGCTCCAACTCACAGTCTCGGGGGATTAAACA
GCATCGTCCAGCAACAGTCAAATCTCAATTCTTAGAGCAATCGAAGCGCTATCTTAAAGACCAAACAGCAGCTTCTTGGCCTCTGGGCCTCTGGGGGATTAAACA
GCTTCAAGCCCGCCGTGCTTGCTATGGAATAGTTCTTGGAGTAATAAATCACAGACCGATÁTTTGGGACAACATGACCTGGTGTAGTGGAAAACTCATC
TGCCCCCACCACCGTGCCTGGGGTTGTTGGGACAACATGACCTGGTGTAGTGGAAAACTCATC
ATAGGGAAATTTCTAATTATACTGGCACAATCTACAAACTCTTGGAAGAAGTCAAAATCAGCAAGAAAGTCAAAAATCAGCAAGAAAGTCAAATGGG
CCTCGCCCCTGGACTCCCTGGAAGAATCTTTGGAGCTGGTTCGACATAACTAATTGGCTGTGTGGtaaagatcttacaa
```

Fig. 37A

Wild-type subtype CRF01_AE

97CNGX2F-AE (854 a.a.)

MRVKETQMNWPNLWKWGTLILGLVIICSASDNLWVTVYYGVPVWRDADTTLFCASDAKAHETEVHNVWATHACVPTDPNPQEIHL
ENVTENFNMWRNNMVEQMQEDVISLWDQSLKPCVKLTPLCVTLNCTNANWTNSNNTTNGPNKIGNITDEVKNCTFNMTTELKDKK
QKVHALFYKLDIVQINSSEYRLINCNTSVIKQACPKISFDPIPIHYCTPAGYAILKCNDKNFNGTGPCKNVSSVQCTHGIKPVVS
TQLLLNGSLAEEEIIIRSENLTNNAKTIIVHLNKSVEINCTRPSNNTRTSITMGPGQVFYRTGDIIGDIRKAYCEINGIKWNEVL
VQVTGKLKEHFNKTIIFQPPSGGDLEIITHHFSCRGEFFYCNTTKLFNNTCIGNTSMEGCNNTIILPCKIKQIINMWQGVGQAMY
APPISGRINCVSNITGILLTRDGGADNNTTNETFRPGGGNIKDNWRSELYKYKVVEIEPLGIAPTRAKRRVVEREKRAVGIGAMI
FGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLRAIEAQQHLLQLTVWGIKQLQARVLAVERYLKDQKFLGLWGCSGKIIC
TTAVPWNSSWSNKSFEEIWDNMTWIEWEREISNYTSQIYEILTESQNQQDRNEKDLLELDKWASLWNWFDITNWLWYIKIFIIIV
GSLIGLRIIFAVLSIVNRVRQGYSPLSFQTPTHHQREPDRPEEIGEGGGEQSKDRSVRLVSGFLALAWDDLRSLCLFSYHLLRDF
ILIAARTVELLGHSSLKGLRRGWEGLKYLGNLLLLYWGQEIKISAISLLNATAIAVAGWTDRVIEVAQRAWRALLHIPRRIRQGLE
RALL

*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF
design and the "W" underlined with red color is the last amino acid at the C
terminus, and all the remaining amino acids after the "W" will be deleted in 140CF
design.

Fig. 37B

97CNGX2F-AE 140CF.pep (629 a.a.)

Nick name: 018

MRVKETQMNWPNLWKWGTLILGLVIICSASDNLWVTVYYGVPVWRDADTTLFCASDAKAHETEVHNVWATHACVPTDPNPQEIHL
ENVTENFNMWRNNMVEQMQEDVISLWDQSLKPCVKLTPLCVTLNCTNANWTNSNNTTNGPNKIGNITDEVKNCTFNMTTELKDKK
QKVHALFYKLDIVQINSSEYRLINCNTSVIKQACPKISFDPIPIHYCTPAGYAILKCNDKNFNGTGPCKNVSSVQCTHGIKPVVS
TQLLLNGSLAEEEIIIRSENLTNNAKTIIVHLNKSVEINCTRPSNNTRTSITMGPGQVFYRTGDIIGDIRKAYCEINGIKWNEVL
VQVTGKLKEHFNKTIIFQPPSGGDLEIITHHFSCRGEFFYCNTTKLFNNTCIGNTSMEGCNNTIILPCKIKQIINMWQGVGQAMY
APPISGRINCVSNITGILLTRDGGADNNTTNETFRPGGGNIKDNWRSELYKYKVVEIEPLGIAPTRARTLTVQARQLLSGIVQQQ
SNLLRAIEAQQHLLQLTVWGIKQLQARVLAVERYLKDQKFLGLWGCSGKIICTTAVPWNSSWSNKSFEEIWDNMTWIEWEREISN
YTSQIYEILTESQNQQDRNEKDLLELDKWASLWNW*

*Amino acids seen in blue color is for easy identification of the junction of the
deleted fusion cleavage site.

Fig. 37C

CODON-OPTIMIZED 97CNGX2F-AE 140CF.seq (1921 nt.)

Nick name: 018

ttcagtcgacagccaccATGGCGAGCAGTAAAAGAGAGACACACAAATGATGAATTGGCCCAATTTGTGGAAGTGGGAACATTGATCCTGGGACT
GGTGATAATCTGTAGTGCGACGCCACAAATCTCTGGGTGACCGTTTACTATGGTGTACCAGTTGGAGAGACGCTGATACCACCCTC
TTCTGTGCAAGCGACGCCAAAGCCCACGAAACCTGAAGTCCATAATGTATGGGCCACCACGCTGCCTACCAACCGACCCTAATC
CCCAAGAGATCCACCTTGAGAATGTAACTGAGAATTTTAACATGTGGAGAAATAACATGGTGGAACAAATGCAGGAAGACGTTAT
TTCCTTGTGGGACCAGAGCCTTAAACCTTGTGTGTCAAATTGACTCCCCTGTGTGTGACTCTCAATTGTACAAACGCAAATTGGACC
AACAGCAACAACAACTACCACGGCCCCTAACAAAATTGGCAATATTACTGCCAATATTACTGATGAAGTCAAGAACTGCACTTTTAACATGACAACAG
AACTGAAGGATAAGAAACAGAAAGTCCATGCTCTGTTCTATAAGCTCGACATAGTACAAATTAATAGCTCAGAATATAGACTGAT
AAACTGCAATACTTCCGTTATCAACGATAAGAATTTTAACGGCACAGGTCCCTGCAAAAACGTTTCCTCTGTCCAGTGTACACACGGTA
TACGCTATCCTGAAATGCAACGATATAGTGCACCTGAGTATATCAACACAACTGTAGAAATCAACTGTACCCGACCCTCAAACAACAACTGAACAAGTATA
TCAAGCCTGTAGTATCAACAACTGCTCCTGAATGGCTCCTTGGCCGAAGAGATCATCATTAGAAGTGAGAACCTGACGAA
CAACGCCAAGACTATAATAGTGCAAGACTATAATCTGTAGAAATCAACTGTACCCGACCCTCAAACAACAACTGAACAAGTATA
ACAATGGGAACGTCCTGGCCAAGTTTTTTACCGGACCGGCGCAACATTTAATAAGAACACATTTAATAAGAACCATAATATTCCAGCCCCCGAGTGGCGG
AGTGGGAACGAAGTACTGGTTCAAGTAACTGGAAAACTCAAAGAAACTCAAAGAAGTCCCTGTAAGATCAAGCAGATTATCAACACGTGC
CGACCTCGAGATTATCACCCCATCACTTTCTTGTAGAGGCGAATTTTTTACTGTAACACGAACAGATCAAGCAGATTATCAACACATGTGGCAGGGAG
ATCGGGAACACTTCTATGGAAGGATGTAATAATACCATTATACTGCCCTGTAAGATCAAGCAGATTATCAACACATGTGGCAGGGAG
TAGGTCAGGCAATGTACGCACCGATTTCAGGACGGATCAATTGCGTATCAAAATATCACCGGCCATTCTGCTGACCCGGGACGG
AGGCGCAGACAACAATACCACTAACGAGACATTTAGACCTGGAGGCGGCAATATAAAGGATAATTGGAGAAGTGAGCTGTATAAA
TACAAAGTCGTAGAGATCGAACCCCTCGGCATTGCTCCAACCCGGACTCTCCAACCTGCTCCAGCTACAAGCTAGACACAGCTGCTTTCTG
GCATAGTCCAACACAGCAGTCAACCTCCTCCGCGCTATTGAAGCAACAACACCTGCTCCAGCTTTGGGGCTGCGTGTGGGAATCAAACA
ATTGCAAGCAAGAGTGCTCGCCTTGGAACTCATCCTGGAGTAATAAAGATCCAAATACTGACAGAAAATACTGGATGATTGAGTGGG
TGTACAACAGCGGTGCCTTGGAACTATACAAGCCAAATTACGAAAATACTGACAGAAAGTCAAAAACCAGCAGGACAGAAATGAGAAAGACCT
GCTCGAACTGGATAAGTGGGCCCTTCTTTGTGGAACTGGTaaagatcttacaa

Fig. 38A

Wild-type DRCBL-G (854 a.a.)

MRVKGIQRNWQHLWNWGILILGLVIICSAEKLWVTVYYGVPVWEDANAPLFCASDAKAHSTESHNIWATHACVPTDPSPQEINMR
NVTENFNMWKNNMVEQMHEDIISLWDESLKPCVKLTPLCVTLNCTEINNSTRNITEEYRMTNCSFNMTTELRDKKAEYALFYR
STQLLNGSLAEKDIIISSENISDNAKVIIVHLNRSVEINCTRPNNNTRRSVAIGPGQAFYTTGEVIGDIRKAHCNVSWTKWNET
LRDVQAKLQEYFINKSIEFNSSSGGDLEITTHSFNCGGEFFYCNTSGLFNNSILKSNISENNDTITLNCKIKQIVRMWQRVGQAM
YAPPIAGNITCRSNITGLILTRDGGDNNSTSEIFRPGGGDMKNNWRSELYKYKTVKIKSLGIAPTRARRRVVEREKRAVGVGAIF
LGFLGTAGSTMGAASITLTVQVRQLLSGIVQQQSNLLRAIEAQQHLLQLTVWGIKQLQARVLALERYLKDQQLLGIWGCSGKLIC
TTNVPWNTSWSNKSYNEIWENMTWIEWEREIDNYTYHIYSLIEQSQIQQEKNEQDLLALDQWASLWSWFSISNWLWYIRIFVMIV
GGLIGLRIVFAVLSIVNRVRQGYSPLSFQTLLHHQREPDRPAGIEGGGEQDRDRSIRLVSGFLALAWDDLRSLCLFSYHRLRDF
ILIAARTVELLGRNSLKGLRLGWEALKYLWNLLLYWARELKNSAINLLDTIAIAVANWTDRVIEVAQGRAVLNIPRRIRQGLE
RALL

*Amino acid sequence underlined is the fusion domain that will be deleted in 140CF
design and the "W" underlined with red color is the last amino acid at the C
terminus, and all the remaining amino acids after the "W" will be deleted in 140CF
design.

Fig. 38B

DRCBL-G 140CF.pep (630 a.a.)
Nick name: 017

MRVKGIQRNWQHLWNWGILILGLVIICSAEKLWVTVYYGVPVWEDANAPLFCASDAKAHSTESHNIWATHACVPTDPSPQEINMR
NVTENFNMWKNNMVEQMHEDIISLWDESLKPCVKLTPLCVTLNCTEINNSTRNITEEYRMTNCSFNMTTELRDKKAEYALFYR
STQLLNGSLAEKDIIISSENISDNAKVIIVHLNRSVEINCTRPNNNTRRSVAIGPGQAFYTTGEVIGDIRKAHCNVSWTKWNET
LRDVQAKLQEYFINKSIEFNSSSGGDLEITTHSFNCGGEFFYCNTSGLFNNSILKSNISENNDTITLNCKIKQIVRMWQRVGQAM
YAPPIAGNITCRSNITGLILTRDGGDNNSTSEIFRPGGGDMKNNWRSELYKYKTVKIKSLGIAPTRARTLTVQVRQLLSGIVQQQ
SNLLRAIEAQQHLLQLTVWGIKQLRARVLALERYLKDQQLLGIWGCSGKLICTTNVPWNTSWSNKSYNEIWENMTWIEWEREIDN
YTYHIYSLIEQSQIQQEKNEQDLLALDQWASLWSW*

*Amino acids seen in blue color is for easy identification of the junction of the
deleted fusion cleavage site.

Fig. 38C

CODON-OPTIMIZED DRCBL-G 140CF.seq (1921 nt.)

Nick name: 017

```
ttcagtcgacagccaccATGAGAGTTAAAGGAATCCAACGCAATTGGCAACACCTTTGGAACTGGGGCATATTGATTCTTGGACT
GGTGATAATTTGTAGCGCTGAAAAACTCTGGGTAACTCTTATTACGGCGTGCCTGTCTCGGGAGGATGCCAACGCCCCCTGTTC
TGCGCAAGTGATGCAAAGGCTCACAGCACTGAATCTCACAACATTTGGGCCACGCCTGTGTGCCAACCGACCCTAGTCCTC
AGGAGATCAACATGAGAAACGTTACCGAAGAAATTTAATATGTGGAAGAATAATATGGTGGAGCAAATGCACGAAGACATAATTTC
ACTCTGGGACGACGAGTCTCTGAAACCATGTGTGAAACTTACCCCCCTGTGCGTCACCCTGAAATCAACAATAACTCA
ACGAGAAATATCACAGAAGAATACCGAATGACTAACCGATTGTTCCTTTAATATGACAACGAACTGCGAGACAAAAGAAGGCTGAAT
ACGCACTTTTCTACCGAACAGATGTTGTACCAATCAACGAGATGAACAATGAAAACAATCCTACCTGGTATAGACT
GACAAACTGTAACGTTAGCACAGTGTCGTGATAAGCAGGCCTGCCCTAAAGTCACATTCGAACCAATACCAATTCACTACTGCGCACCCGCC
GGATTCGCTATTCTTAAGTGCGTGGATAAGAAGTTAACGGAACTGGAACCTGCAATAATGTATCTACAGTACAATGCACGCATG
GAATTAAGCCTGTCGTTCAACCCCAGTTGCTGCCTCAAACCGCTCAGTTGAAATAAACTGCACTCGGCCTAATAATAACAAGACGCTCT
TGATAATGCAAAGGTCATCATCGTCCACCTGAAAGTCCAGTTGAAATAAACTGCACTCGGCCTAATAATAACAAGACGCTCT
GTCGCAATCGGCCCAGGACAAGCTTTTTACACTACCGGGAAGTTATCGGCGACATACGAAAGCCCACTGCAACGTTAGCTGGA
CCAAGTGGAATGAAAACACTGCGCGATGTTCAAGCCAAACTTCATAAACAAATCATTGAGTTCAATTCAGCTC
TGGCGGCGACCTCGAGATTACAACTCACTCCTTTAACTGCGGCGAATTCTTTATTGTAATACCTCCGTCTCTTCAACAAC
TCTATCCTCAAAGTAACATTTCTGAAAATAATGACACAACTGAATTGCAAGATCAAGCAGATTGTTGATGTGGCAAC
GAGTCGGCAAGCTATGTACGCCCCATCGCCGGAAATATAACGTGTCGATCAAAATATACCTGGCCTCATCCTTACTAGAGA
TGGCGGAGACAATAATAGCACCAGAGATATTCAGACCAGGCGGAGATGAAAAAACAACTGGAGGTCCGACGTCTCACAAG
TACAAAACAGTCCAACAACAGTCAAAATTAAAGCCTGGCATTGCTCGCGCCTATAGAGGCTCAACAACAACAACTCCTGGGCATATGGGAAAACCATGGGGTGTTCCGGCAAACTGATC
GAATCGTCCAACAACAGTCCAACTTGCTCGCGCCTATAGAGGCTCAACAACAACAACTCCTGGGCATATGGGAAAACCATGGGGTGTTCCGGCAAACTGATC
ATTGAGAGCAAGAAATTGACAATTATACCATATACCATATACTCTCTCATCGAACAATCTCGAACAATCTCTCCATCGAACAATCTCGAACAATCTCTCTCATCGAACAATCTCTCGAACAATCTCTCTCATC
TGCACAACAAATGTACCCTGGAACACCAGCTGCTCGGAACGGTATCTTAAGGACCAACAACAACTGGTGTTCCGGCAAACTGATC
AAAGGGAAATTGACAATTATACCATATACCATATACTCTCTCATCGAACAATCTCTGAACAATCCTGAACAATCTCTCATCGGGG
GTTGGCTCTTGACCAATGGGCTTCTTTGTGGAGTTGGtaaagatcttacaa
```

SUBSTITUTE SHEET (RULE 26)

2003 Centralized HIV-1 Envelope Proteins and the Codon-Optimized Gene sequences

Fig. 39A

2003 Cons Env

MRVMGIQRNCQHLWRWGILIFGMLIICSAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEIVLENVTENF
NMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVNATNNTTNNEEIKNCSFNITTEIRDKKKKVYALFYKLDVVPIDDNNSYRLI
NCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIRSENITNNAKTIIV
QLNESVEINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNISRTKWNKTLQQVAKKLREHFNKTIIFNPSSGGDLEITTHSFNCGGE
FFYCNTSELFNSTWNGTNNTITLPCRIKQIINMWQGVGQAMYAPPIEGKIRCTSNITGLLLTRDGGNNNTETFRPGGGDMRDNWRSELYKYK
VVKIEPLGVAPTKAKRRVVEREKRAVGIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLRAIEAQQHLLQLTVWGIKQLQAR
VLAVERYLKDQQLLGIWGCSGKLICTTNVPWNSSWSNKSQDEIWDNMTWMEWDKEINNYTDIIYSLIEESQNQQEKNEQELLALDKWASLWN
WFDITNWLWYIKIFIMIVGGLIGLRIVFAVLSIVNRVRQGYSPLSFQTLIPNPRGPDRPEGIEEGGGEQDRDRSIRLVNGFLALNIPRRIRQGFERAL
L$

Fig. 40A

2003 M. Group.AnC Env

MRVMGIQRNCQHLWRWGILIFGMLMICSAAENLWVTVYYGVPVWKEANTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEIVLENVTENI
NMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVNATNNSTNMGEIKNCSFNITTEIRDKKQKVALFYRLDVVPINDNNSYRLJ
NCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIRSENITDNAKTII\
QLNESVEINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNISGAEWNKTLQQVAAKLREHFNNKTIIFKPSSGGDLEITTHSFNCGG
EFFYCNTSGLFNSTWNGTNETITLPCRIKQIVNMWQRVGQAMYAPPIAGNITCKSNITGLLLTRDGGTNNTETFRPGGGDMRDNWRSELYKY
KVVKIEPLGVAPTKAKRRVVEREKRAVGIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLRAIEAQQHLLQLTVWGIKQLQA
RVLAVERYLKDQQLLGIWGCSGKLICTTNVPWNSSWSNKSQDEIWDNMTWMQWEREISNYTDIIYSLIEESQNQQEKNEQDLLALDKWASLW
NWFDITNWLWYIKIFIMIVGGLIGLRIVFAVLSIVNRVRQGYSPLSFQTLIPNPRGPDRPGGIEEGGEQDRDRSIRLVSGFLALAWDDLRS
LCLFSYHRLRDFILIAARTVELLGRRGWEALKYLMNLLQYWGQELKNSAISLLDTTAIAVAEGTDRVIEVVQRACRAILHIPRRIRQGFERA
LL$

SUBSTITUTE SHEET (RULE 26)

Fig. 40B

```
2003 M. Group.anc Env.seq.opt
ATGCGCGTGATGGGCATCCAGCGCAACTGCCAGCCACCTGTGCGCGCTGGGCATCCTGATCTTCGGCATGCTGATGATCTGCTCCGCGCCCGA
GAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGTGAAGGAGGCCAACACCACCTGTTCTGCGCCTCCGACGCCAAGGCCTACGACA
CCGAGGTGCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTCCTGGGGAACGTGCTGACCGAGAACTTC
AACATGTGGAAGAACAACATGGTGGAAGAACTCACCAACAACTCCCTGTGGGACCCAGTCCTGCGTGAAGCCCTGCGTGCTGAAGCTGACCCC
CCTGTGCGTGACCCTGAACTGCAACGACGCGAAACTGCACCGACGTGAACGCGTGTGGGCGCCCATCAAGGGCGAGATCAAGAACTGCTCCTTCAACATCACCA
CCGAGATCCGCGACAAGAAGCAGAAGGTGTACGCCCTGTTCTACCGCCTGGACGTGGTGCCCATCAACGACACAACTCCTACCGCCTGATC
AACTGCAACACCTCCGCCATCACCCAGGCCTGCCCCAAGGTGTCCTTGCAAGAACGTGCACCGTGCACCGACTACTGCGCCCCCGGCTTCGCCAT
CCTGAAGTGCAACGACAAGAAGTTCAACGGCGTCCCTGGCCGAGGAGGAGATCATCATCCGCAGTGTCCACCCGCTCCGAGAACATCACCGACAACGCCAAGACCATCATCGTG
CAGCTGAACGAGTCCGTGGAGATCAACTGCACCCGCGCCCACTGCAACATCTCCGGCGCCGCACGCCCTGCGCCCAGCCAGGTGGCCGCCAAGC
CACCGGCGACATCATCGGCGACATTCAACAACAAGACCATCATCTTCAAGCCCTCCTCCGGCGGCGACCTGGAGATCACCACCCACTCCTTCAACTGCGGCGGC
TGCCGCGAGCACTTCAACAACAAGCAGACCATCATCTTCAAGGCCTCCGGCTGTTCTTCTACTGCAACACCTCCGCGCCTGTTCAACTCCAACTCCGCCATCAAGCAGAT
CGTGAACATGTGGCAGCGCGTGGGGCCAGCCATGTACGCCCCCCATCCTGCCGCGGCGGCCACCAACAACAACACCGAGACCTTCCGCCCCACCGGCGGCGGCGACATGCGCGAC
TGACCCGCCGACGACGGCACCAACAACAACAACCTGGGCGTGGCCCCCACCAAGGCCAAGCGGCCGAGAAGGCGCGAGCATCGGGCGC
AAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCGGCTCCACCATGGGCGCCCTGCTGTGTGGCCCGGCCATCGGGCGC
CGTGTTCCTGCAGCAGCAGTCCAACCTGCTGCGCGCCATCGAGGCCCAGCACCTGCTGCAGCTGCGCGCCGTGTGGGGCATCAAGCAGCTGCAGGCC
TCGTGCCGCGTGGCCGTGCCCGTGCTACCTGAAGGACCAGCAGCTGCTGGGCGTCTGGGGCTGCTCCGGCAAGCTGATCTGCACCACCCAAGTGCCC
CTGGAACTCCTCCTGGTCCAACAAGTCCCAGGAGGAGTCCCAGACCAGCAGGAAGAAGAACCAGCAGGAGGACCTGCTGGCACTGGACAAGTGGGCCTCCCTGTGG
ACATCATCTACTCCCTGATCGAGGAGTCGTGTTGTACATCAAGAATCTTCATCATGATCGTGGGCGGCCTGATCGGCCTGCGTATCGTGCTGCCCGT
AACTGGTTCGACATCACCAACTGGCTGTGGTACATCAAGCTACTCCCCCTGTCCTTCCAGACCCTGTCCCGCCTTCCTGGCCTCCGTGTCCCGCACCGTGTCCCGCACCGTGGGCG
GCTGTCCATCGTGAACCGCGTGCGCCAGGGCTACTCCCCGCTGTCCTTCCAGCAGACCCTGCTGCCGGCTCCTC
GCATCGAGGAGGAGGGCGGCGAGCAGGACCGCGATCGAGGGCGGCGAGCAGGACGGCACCGGCGCCATCGCGAGCTGCTGGGACGCTCCGCTCC
CTGTGCCTGTTCTCCTACCACCGCCTCGCGCGACCTTCATCCTGATCGCCGTGCGCCTCGCTGATCGGGGGCCAGGAGCTGAAGAACTCCGCCATCCTCCCCTGTGCCCGCACCGTGTCCCGTCCC
GAAGTACCTGTGGAACCTGCTGCAGTACTGGGGCCAGGAGCTGAAGAACTCCGCCATCAGCCTGCTGGGACGCCTGCTGGGACGCCTGGGAGGCCCT
AGGGCACCGACCGCGGCGTGGTGCAGCGCGTGATCGAGTGGCTGCAGACCGCCATCCTGCACGCCATCCCGCGCCGCATCCGCCAGGGCTTCGAGCGCGCGCC
CTGCTGTAA
```

Fig. 41A

2003 CON A1 Env

```
MRVMGIQRNCQHLLRWGTMILGMIIICSAAENLWVTVYYGVPVWKDAETTLFCASDAKAYETEMHNVWATHACVPTDPNPQEIHLENVTEEF
NMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLNCSNVNVTNNTTNTHEEEIKNCSFNMTTELRDKKQKVYSLFYRLDVVQINENNSNS
SYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKEFNGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEVIIRSENITNNA
KTIIVQLTKPVKINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSRSEWNKTLQKVAKQLRKYFKNKTIIFTNSSGGDLEITTHS
FNCGGEFFYCNTSGLFNSTWNNGTMKNTITLPCRIKQIINMWQRAGQAMYAPPIQGVIRCESNITGLLLTRDGGNNNTNETFRPGGGDMRDN
WRSELYKYKVVKIEPLGVAPTRAKRRVVEREKRAVGIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLRAIEAQQHLLKLTV
WGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTNVPWNSSWSNKSQNEIWDNMTWLQWDKEISNYTHIIYNLIEESQNQQEKNEQDLLA
LDKWANLWNWEDISNWLWYIKIFIMIVGGLIGLRIVFAVLSVINRVRQGYSPLSFQTHTPNPRGLDRPGRIEEEGEQGRDRSIRLVSGFLA
LAWDDLRSLCLFSYHRLRDFLLIAARTVELLGHSSLKGLRLGWEGLKYLWNLLLYWGRELKISAINLVDTIAIAVAGWTDRVIEIGQRIGRA
ILHIPRRIRQGLERALL$
```

Fig. 42A

2003 A1.AnC Env

```
MRVMGIQRNCQHLLMRWGTMIFGMIIICSAAENLWVTVYYGVPVWKDAETTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEIDLENVTEEF
NMWKNNMVEQMHADIISLWDQSLKPCVKLTPLCVTLNCSNVNVTNNTTNTHEEEIKNCSFNMTTELRDKKQKVYSLFYRLDVVPINENNSNS
SYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKEFNGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEVMIRSENITDNA
KTIIVQLTEPVKINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSRTEWNKTLQKVAAQLRKHFNNKTIIFNSSSGGDLEITTHS
FNCGGEFFYCNTSGLFNSTWNNGTMKDTITLPCRIKQIINMWQRVGQAMYAPPIQGVIRCESNITGLLLTRDGGNNNTNETFRPGGGDMRDN
WRSELYKYKVVKIEPLGVAPTRAKRRVVEREKRAVGLGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLRAIEAQQHLLKLTV
WGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTNVPWNSSWSNKSQDEIWDNMTWLQWDKEISNYTDIIYNLIEESQNQQEKNEQDLLA
LDKWANLWNWFDISNWLWYIKIFIMIVGGLIGLRIVFAVLSVINRVRQGYSPLSFQTLTPNPEGPDRPGRIEEEGEQGRDRSIRLVSGFLA
LAWDDLRSLCLFSYHRLRDFLLIAARTVELLGRSSLKGLRLGWEGLKYLMWNLLLYWGRELKISAINLLDTIAIAVAGWTDRVIEIGQRICRA
ILNIPRRIRQGLERALL$
```

Fig. 41B

2003 CON A1 Env.seq.opt

```
ATGCGCGTGATGGGCATCCAGCGCCAACTGCTGGGCGCTGGGGCACCATGATCCTGGGCATGATCATCTGCTCCGCGCCGA
GAACCTGTGGGTGACCGTGTACTACGGCGTGCCCCTGGAAGGACGCCGAGACCACCCCGACCCCCAGACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGAGA
CCGAGATGCACAACACGTGTGGGCCACCCACGCGCCTGCCGTGCCACCGACATCATCTCCCTGTGGGACCAGTCCTGTGCCGTGCGTGAAGCCTGGAAGCTGACCCC
AACATGTGCGAAGAACAACATGGTGGAGCAGATGCACAACGTGACCAACGTGTTCTCCTGGAAGGTGTACTCACCCAGCACGGCTGTGGGCCACCGACGCTCCAACTCC
CCTGTGCCTGACCCTGAACTGCTCCAACGTGACCGTGAAGGTGTACTCCCTGTTCTACCGCCTGGACGTGGTGCAGATCAACGAACAACTCCAACTCC
TGACCACCGAGCTGCGCGACAAGAAGCAGAAGGTGTACTCCCTGTTCTACCGCCTGGACGTGGTGCAGATCAACGAACAACTCCAACTCC
TCCTACCGCCTGATCAACTGCAACACCTCCGCCATCACCCAGGCCTGCCCCAAGGTGTCCTTCGAGCCCATCCCCATCACTGCGCCC
CGCCGGCTTCGCCATCCTGAAGTGCAAGGACAAGGAGTTCAACGGCACCGGCCCCTGCAAGAACGTGTCCACCGTGCAGTGCACCCACGGCA
TCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGGAGGTGATCATCCGCTCCGAGAACAACAACACCCCGG
AAGACCATCATCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGGAGGTGATCATCCGCTCCGAGA
CCAGGCCTTCTACGCCGTGGCAGCTGCACCGGCGACATCATCGGCCAAGTGCTGCCAGGTGGAACAAGAACCCTGCAGA
AGTGGCCAAGCAGCTGCAGCGAGTTCTTCTACTGCAACAACACCTCCGGCCTGGCTTCAACCTGGCAGTGGCCGTGAGT
TTCAACTGCGGCGCATCAAGCAGCATCAACGATCAACGTGGCCAGCGGCGCCCCCATCGAGGCGTTGATCCGCTGCGAGT
GCCCTGCCCGCATCAAGCAGCATCAACGATCAACGTGGCCAGCGGCGCCAACAACACCAGGCGCCCCATCGAGGCGTTGATCCGCTGCGAGT
CCAACATCACCGGCCTGCTGCTGACCCGCGACGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCAAGCGCCGGCGCCCGTGGTGGAGCGCGA
TGGCCGGCTCGTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCAAGCGCCGGCGCCCGTGGTGGAGCGCGA
GAAGCGCGCCGTGGGCATCGGCGCCGTGTTCCTGGGCTTCCTGGGCGCCGCGGGCTCCACCATGGGCGCCGCCAGCATGACCCTGACCGTG
AGCGCCCGGGCCGTGCAGCTGCTGTCGGGCATCGTGCAGCAGCAGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGAAGCTGACCGTG
TGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAACGTGCTGGCCGTGGAACTCCTCCGTGCCGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAACGTGCTGGCCGTGGAACTCCTCCGTGCCGGGCATCAAGCAGCTGCAG
GCTGATCTGCACCACCACTGCCGTGCCCTGGAACTCCATCTCCAACATCATCTACACAACTGGTTCGACATCTCCAACTGGTTCGACATCTCCAACTGGTTCGACATCTCCAACTGGTTCGACATCTCCAACTGGTTCGACATCTCCAAC
ACAAGGAGATCTCCAACTACACACCCAACTGTGGAACTGGTTCGACATCTCCAACTGGTACATGACCTGGCTGCAGTGGGATAAGGAGATCTTCATCATGAT
CTGGACAAGTGGGCCAACCTGTGGAACTGGTTCGACATCTCCAACTGGTACATGACCTGGCTGCAGTGGGATAAGGAGATCTTCATCATGAT
CGGCCTGCTGGCCATCGTGTTCGGGCCATCGTGCTGGGCAACCTGTTCCTGCCCTGGGAGCTGGAACTCCGTGGACTGGCTTCCTGGCC
CCGGCGGCCTGGACCGCCCGGCCATCGTGCTGGGCAACCTGTTCCTGCCCTGGGAGCTGGAACTCCGTGGTGGGCGGCTTCCTGGCC
CTGGCCTGGGACGACCTGCGCTCCCTGTGCCTGTTCTCCTACCACCGCGTGGGCATCTCCCTGGAACCTGGAGGCCTGGAGCTGCTGGTGCTGGGGCTGGAGCTGCT
GGGCCTGCGCATCCTGGAAGGCCTGCGCCTGGGCTGGGGAGGGCCTGGGCCATCGTGCTGGGCGCCTGAAGTACCTGGGCAACCTGCTGAGCTACTGG
TCTCCGCCATCAAGGCCTGGTGGACACCATCCGCGCCGTGGACCGGCTGGCCGGCTGATCGAGATCGGCGCCCAGCGCCCATCCGCCGCGCC
ATCCTGCACATCCCCCGCCGCATCCGCCAGGGCGCCGCCCCTGCTGTAA
```

Fig. 42B

2003 A1.anc Env.seq.opt

ATGGCGGTGATGGGCATCCAGCGCAACTGCCAGCCACCTGTGGCGCTGGGGCACCATGATCTTCGGCATGATCATCATCTGCTCCGCGCCGA
GAACCTGTGGGTGACCGTGTACTACGGCGTGCCGGTGTGGAAGGACGCCGAGACCACCACCCTTCGCGCCTCCGACGCGCCAAGGCCTACGACA
CCGAGGTGCACAACGTGTGGGCCACCACGTGCGCCGACCCCAACCCCAGGAGGATCGACCTGGAGAACGTGACCGAGGAGTTC
AACATGTGGAAGAACAACATGGTGAGCAGATGCACGCCGACATCATCTCCGTGGGGACCAGTCCCTGAAGCCTGAAGCTGACCCC
CCTGTGCGTGACCCTGAACTGCTCCAACGTGACCGACGAAGGTGTACTCCCTGTTCTACCGCCTGGTCCATCACCCCACAGGTCGCTCCTTCAACA
TGACCACCGAGCTGCGCGACAAGAAGCAGAAGGTGTACTCCCTGTTCTACCGCCTGGTCCATCACCGACGAACAACTCCAACTCC
TCCTACCGCCTGATCAACTCCGCCATCCTGGACGCTGGTGCCCATCAACCCGACCTTCGCCATCCCCATTCCACTACTGCGCCC
CGCCGGCTTCGCCATCCTGGGTGTCCACCCGTGGGTGGTGCCCAAGGTGTCTTCGCAAGAACGTGTCCGCTCCACTGCACCCACGGCA
TCAAGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGCTCCCTGGCCGGAGGAGGTGATGATCCGCTCCGAGAACATCACCGACAACGCC
AAGACCATCATCGTGCAGCTGAACGGCCGTGAAGATCAACTGCACCCGCCCCAACAACAACACCCGCAAGTCCATCCGCATCGGCCCGGA
CCAGCCTTCTACGCCCAGCTGGCCAACATCGGCGACATCATCCGCCAGCCCACCGCCAACGTGTCCCGGCGGACCTGGAACAAGACCCTGCGAGA
AGTTGGCCGCCCAGCTGCGCAAGCACTTCAACAACAAGACCATCATCTTCAACCTCCTCCTCCGGGAACAACGGCCACCTGGAGATCACCACCACTCC
TTCAACTGGCGGCGGAGTTCTTCTACTGCAACATCATCAACATGTGGCAGCGCGTGGGCCAGGCCGTGATCCGCTGCGAGT
GCCCTGCCGCATCAAGCACAGATCATCAACAACATGTGGCAGCGCGTGGGCCAGGCCGTGATCCGCTGCGAGT
CCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAACAACAACAACACCCACCTTCCGCCCCGGCGGCGGCCAAGCGCCGCCCTGGTGGGAGCGCGA
TGGCCGCTCCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCGCGTTCCTGGGCGTGTTCCTGGGCGTCGTCCAACCTGCTGGCCGTGTCGCCCCGGCGGAGGCGCGA
GAAGCGCCGCCCGGCCGTGGGGCGTGTTCCTGGGGCCGTCGTCTGCAGCAGCAGTCCAACCTGCTGGCCGTGTCGCCCGTTCCTGGGCGTCGTCCAACCTGCTG
AGGCCCGCCAGCTGCTGTCCGCGCCTGGGGCCGTCGTCTGCAGCAGCAGTCCGCGTGCTGGCCGTGTGCCGCGCATCGTGCTCCATCACCCTGACCGTG
TGGGGCATCAAGCAGCTGCAGGCCGGCCAGCGCCTGCTGCGCGCCGTGGGCCGTCCTGAGCCTGTGCTGTCCAACCTGCTGGCCGTGTTCCTGAGCCTGTCGCCGTGCTGCGCGCATCGTGCTCCATCACCCTGACCGTG
GCTGATCTGCACCACCAACGTGCCCTGGAACTCCTCCTGGTCCAACATCTCACCGACAAGAGTCCCAGGACGAGAGATCTCCAGAAGATCTTGGGACAACAGCAGGAGAAGAACAGCAGAGATCTTCATCATGATCGTGGGCGGCCTGAT
ACAAGGAGATCTGCCACCAACGTGCCCTGGAACTCCTCCTGGTCCAACATCTCACCGACAAGAGTCCCAGGACGAGAGATCTCCAGAAGATCTTGGGACAACAGCAGGAGAAGAACAGCAGAGATCTTCATCATGATCGTGGGCGGCCTGAT
CTGGACAAGTGGGCCAACCTGTGGAACTGGTTCGACATCTCCAACTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGAT
CGGCCTGCGCATCGTGTTCGCCGTGCTGTCCGTGATCAACCGCGTGCGCCAGGGCTACTCCCCGCTGTCCTTCCAGACCCCTGACCCCCAACC
CCGAGGGGCCCCGACCGAGCCCTCCCCGTGCGCGGGGGCCGCGCTGGGCGTGGGAGGGCCTACCCACCGCCGTGTCCGGCTTCCTGGCC
CTGGCCTGGGACGACGACCTGCCTGAAGGGCTCCGTGTTCTCCGTGTTCCTGTGCGGGGGCGGCCTGGGCGTGGGAGGGCCTACCCACCGCCGTGTCCGGCTTCCTGGCC
GGGGCCGCGTCCTCCGCGATCAACTGCTGGACACCCGCCATCGCCATCCTGCTGCTCATCCTGATCGCCCCGCGTGGGGCCCGTGCTGCTGCTACTGGGCCGCGAGCTGCT
TCTCCGCCATCAACTGCTGCGCGATCAACTGCTGGACACCCGCCATCGCCATCCTGCTGCTCATCCTGATCGCCCCGCGTGGGGCCCGTGCTGCTGCTACTGGGCCGCGAGCTGCT
ATCCTGAACATCCCCCGCCGCATCCGGCCTGGGGGCCTGGAGCGCGCGCCAGCCAGCCATCTGCCGCGCC
GGGGCCGCGTCCTCCGCGATCAACTGCTGGACACCCGCCATCGCCATCCTGCTGCTCATCCTGATCGCCCCGCGTGGGGCCCGTGCTGCTGCTACTGGGCCGCGAGCTGCT
ATCCTGAACATCCCCCGCCGCATCCGGCCTGGGGGCCTGGAGCGCGCGCCAGCCAGCCATCTGCCGCGCC
TAA

Fig. 39B

```
2003 CON-S Env.seq.opt
ATGCGCGTGATGGGCATCCAGCGCAACTGCCAGCACCTGTGGCGCTGGGCATCCTGATCTTCGGCATGCTGATCATCTGCTCCGCCGCCGA
GAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAACACCACCCTGTTCTGCGCCTCCAAGGCCAAGGCCTACGACA
CCGAGGTGCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCGTGCTGGAGAACGTGACCGAGAACTTC
AACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCC
CCTGTGCGTGACCCTGAACTGCACCGACGTGAACGCCACCAACACCACCAACCAGAGCGATCAAGAGGAGATCAAGAACTGCTCCTTCAACATCACCA
CCGAGATCCGCGACAAGAAGAAGAAGGTGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACGACAACAACTCCTACCGCCTGATC
AACTGCAACACCTCCGCCATCACCCAGGCCTGCCCCAAGGTGTCCTTCGAGCCCATCCCCATCCACTACTGCGCCCCGGCTTCGCCAT
CCTGAAGTGCAACGACAAGAAGTTCAACGGCACCGGCCCCTGCACCAACGTGTCCACCGTGCAGTGCACCCACGGCATCCGCCCCGTGGTGT
CCACCCAGCTGCTGCTGAACGGCAGCCTCCCTGGCCGAGGAGGAGATCATCATCCGCTCCGAGAACATCACCAACAACGCCCCCGCCAGGCCTTCTACGC
CAGCTGAACGAGTCCGTGGAGATCAACTGCACCCGCCCCAACAACAACACCCGCAAGTCCATCCGCATCGGCCCAGGGCCGCATCGGCCCGC
CACCGGCGACATCATCGGCGACATCCGCCAGGCCCGCTGTTCAACTACTTCAAGCCCTCCGGGCCGGCACCAACACCATCACCCTGCCCTGC
TGCGCGGAGCACTTCAACAAGACCATCAAGGTGGAACAAGTGGAAGCTTGGAGATCAACAACAACATCACCACCCTGCCCTGCCGCATCAAGCAGATCAT
TTCTTCTACTGCAACACCTCCGAGCTGTTCAACTCCAACGGCAACGGCACCAACGGTGTCCCCCCATCGAGGGCCAAGATCCGCCTGCTGTGA
CAACATGTGGCAGGGCGTGGGCAAGGCCATGTACGCCCGCCCAGGCCCTGGGCCGGGCGACAACACCGAGGACCTTCCCGGGCGTGGGCCCCGG
CCCCGGGCCGCAACAACACCGAGACCTTCCGCCCCACCAGGGCGTGGGCGCCCAAGGGCCCGCGCCCAGCGCCGTGGGGCATCGGCGCCGGI
GTGGTGAAGATCGAGCGCCTTCCTGGGCGGCCGCCGGCTCCACCATGGGCGCCCGCCTCCAGCACCGTGCAGGCCCCAGCTGCAGGCCCGC
GTTCCTGGGCGTTCCTGGGCGCCGCCGGCTCAACCTGCGCGCCCATCGAGGGCCCAGCAGCTGCTCAAGCAGCTGCAGGCCCCGC
TGCAGCAGCAGTCCAACCTGCTGCGCGCTACCTGAAGGACCAGCAGCTGCTGGGCGTGGGGGCATCAAGCAGCTGACCGTGTGGGCATGATCTGGGGGCATGATCTGCACCAACGTGCCCTG
GTGCTGGCCGTGCCGTGCTACCTGAAGGACGCCTACCAAGAGTCCAGGACAACAAGAGCAAGCTGGACAGCATCTGGGACAACATGACCTGGATGGAGTGGGACAGAGAAGAACTACACCGACA
GAACTCCTCCTGGTCCAACAAGAGTCCCAGGACGAACAACATGACCTGGGATGGAGTGGGACAAGAGGACAAGAACTACACCGACA
TCATCTACTCCCTGATCGAGGAGTCCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGCCCTGGACAAGTGGGCCTCCCTGTGGAAC
TGGTTCGACATCACCAACTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCTTCCTGGGCGTGCTGCGCCTGTTCGCCGTGCT
GTCCATCGTGAACCGCGTGCGCCAGGGCTACTCCCCGCTGTCCCTCCGGGCCACCGCCTCCCTGGGCCTGCTGTGCCCCGCGCCCTGCTCCCTG
TCGAGGAGGAGGGCGGCGAGCAGGACCGCGACCGCAGCATCCGCCTGGTGAACGGCTTCCTGGCCCTGGCCTGGGACGACCTGCGCTCCCTG
TGCCTGTTCTCCTACCACCGCCTGCGCGACCTGCTGCTGATCGTGGGGCGTACTGGGGCCGGCCGGCAGCTGCTGGGCGGCCTGGCCGAGG
GTACCTGTGGAACCTGCTGCAGTACTGGGGCCCAGGAGCTGAAGAAČTCCGCCGTGAGCCTGCTGAACGCCACCGCCATCGCCGTGGCCGAGGGCGACCGGGACCGGCCGGCGGCCGCCGACCGCCCAGGCCGTTGCGGCGCCGGCGGGCGCCCTGCCCGACGCCCCAGCGCCTTCGAGCGCGGCGCCCTG
GCACCGGACCGCGTGATČGAGGTGGTGCAGCGCGCTGCGAGAGGTGGTGCAGCGCGCCAGGCCCCCCGCATCCCGCCGCATCCTGAACATCCCCCGCCGCATCCTGAACATCCCCCGCCGCATCCTGAACATCCCCCGCCGCATCCTGAGCGCCTTCGAGCGCGGCGCCCTG
CTGTAA
```

Fig. 43A

2003 CON_A2 Env

MRVMGTQRNYQHLWRWGILILGMLIMCKATDLMWTVYYGVPVWKDADTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEVNLENVTEDFN
MWKNNMVEQMHEDIISLMDQSLKPCVKLTPLCVTLNCSNANTTNNSTMEEIKNCSYNITTELRDKTQKVYSLFYKLDVVQLDESNKSEYYYR
LINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDPRFNGTGSCNNVSSVQCTHGIKPVASTQLLLNGSLAEGKVMIRSENITNNAKNI
IVQFNKPVPITCIRPNNNTRKSIRFGPGQAFYTNDIIGDIRQAHCNINKTKWNATLQKVAEQLREHFPNKTIIFTNSSGGDLEITTHSFNCG
GEFFYCNTTGLFNSTWKNGTNNTEQMITLPCRIKQIINMWQRVGRAMYAPPIAGVIKCTSNITGIILTRDGGNNETETFRPGGGDMRDNWR
SELYKYKVKIEPLGVAPTRAKRRVVEREKRAVGMGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLKAIEAQQHLLKLTVWG
IKQLQARVLALERYLQDQQLLGIWGCSGKLICATTVPWNSSWSNKTQEEIWNNMTWLQWDKEISNYTNIIYKLLEESQNQQEKNEQDLLALD
KWANLWNWFNITNWLWYIRIFIMIVGGLIGLRIVIAIISVVNRVRQGYSPLSFQIPTPNPEGLDRPGRIEEGGEQGRDRSIRLVSGFLALA
WDDLRSLCLFSYHRLRDCILLAARTVELLGHSSLKGLRLGWEGLKYLWNLLLYWGRELKNSAISLLDTIAVAVAEWTDRVIEIGQRACRAIL
NIPRRIRQGFERALLS

Fig. 44A

2003 CON_B Env

MRVKGIRKNYQHLWRWGTMLLGMLMICSAAEKLWVTVYYGVPVWKEATTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEVVLENVTENF
NMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLMNATNTNTTIIYRWRGEIKNCSFNITTSIRDKVQKEYALFYKLDVVPIDND
NTSYRLISCNTSVITQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAEEVVIRSENFTD
NAKTIIVQLNESVEINCTRPNNNTRKSIHIGPGRAFYTTGEIIGDIRQAHCNISRAKWNNTLKQIVKKLREQFGNKTIVFNQSSGGDPEIVM
HSFNCGGEFFYCNTTQLFNSTWNNTEGNITLPCRIKQIINMWQEVGKAMYAPPIRGQIRCSSNITGLLLTRDGGNNETEIFRPGGGDM
RDNWRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRAVGIGAMFLGFLGAAGSTMGAASMTLTVQARQLLSGIVQQQNNLLRAIEAQQHLLQ
LTVWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTAVPWNASWSNKSLDEIWDNMTWMEWEREIDNYTSLIYTLIEESQNQQEKNEQE
LLELDKWASLWNWFDITNWLWYIKIFIMIVGGLVGLRIVFAVLSIVNRVRQGYSPLSFQTRLPAPRGPDRPEGIEEGGERDRDRSGRLVDG
FLALIWDDLRSLCLFSYHRLRDLLLIVTRIVELLGRRGWEVLKYWWNLLQYWSQELKNSAVSLLNATAIAVAEGTDRVIEVVQRACRAILHI
PRRIRQGLERALLS

SUBSTITUTE SHEET (RULE 26)

Fig. 43B

2003 CON_A2 Env.seq.opt

[DNA sequence figure — nucleotide sequence shown vertically on page]

Fig. 44B

2003 CON B Env.seq.opt
```
ATGGCGCGTGAAGGGCATCCGCAAGAACTACCAGCACCTGTGGCGCTGGGCGCTGTGGGGCACCATGCTGCTGTGGGCATGCTGCCGCGCCGA
GAAGCTGTGGGTGACCGTGTACTACGGCGTGCCGGTGTGGAAGGAGGCCACCACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGACA
CCGAGGTGCACAACGTGTGGGCCACCACGCCCTGCCCGACCCCAACCCCAGAGAGGTGGTGCTGGAGAACGTGACCGAGAACTTC
AACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCACCATCATCACAGAGTCCCTGAAGCTGACCCC
CCTGTGCGTGACCCTGAACTGCACCGACCTGATGAACGCCACCATCATCTACACCGGCTGGCGGCGGCGAGATCAAGAACT
GCTCCTTCAACATCACCACCTCCATCCGCGACAAGGTGCAGAAGGAGTACGCCCTGCTGTACAAGCTGGACGTGGTGCCCATCGACAACGAC
AACACCTCCTACCGCCTGGTTCGCCATCCGTGATCTCCTGCAACGACACCACCTCCGTGACAAGAAGTTCAACGGCCTGCCCAAGGTGTCCTTGCCCAC
CGCCCCCGCGCGGCGTTCGCCATCCGTGATCTCCTGCAACGACACCACCTCCGTGACAAGAAGTTCAACGGCCTGCCCAAGGTGTCCTTCAACGGC
ACGGCATCCGCCCCCGTGGTGTCCCAGCTGCTGCTGAACGGCTCCGTGGCCGAGGAGTGGTGATCCGCTCCGAGAACTTCACCGAC
AACGCCAAGACCATCATCGTGCAGCTGAACGAGTCCGTGGAGATCAACTGCACCCGCCCGAACAACACCCGCAAGTCCATCCACATCGG
CCCCGGCCGCGCCTTCTACACCACCGGCGACAATCGGCGACATCATGGCCGACTGCAACATCTCCCGCGCCGGGCGCACCTGGAACAACACCC
TGAAGCAGATCGTGAAGAAGCTGCGCGAGCAGTTCGGCGCGAGTTCTTCTACTGCAACAGCACCCAGCTCGTGTTCAACTCCACCTGGAACGGCAC
CACTCCTTCAACTGCGGCGGCGAGTTCTTCTACTGCAACAGCACCCAGCTCGTGTTCAACTCCACCTGGAACGGCACCTGGAACACACCGAGGG
CAACATCACCCTGCCCTGCCGCATCAAGCAGATCATCAACATGTGGCAGGAGGTGGCAACAACGAGACCGAGATCTTCCGCGGCGGCGGCGACATG
TCCGGCTGCTCCTCCAACATCACCGGCCTGCTGCTGACCCGCGACCAACTGGCGCTCCGAGCTGTGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCGCCGCGGTGGT
GCAGGCGGAGAAGCGCGCCGTGGGCATCGGCGCCGTGTTCCTGGGCTTCCTGGGAGCAGCAGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGCAG
CTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGTGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTG
CTCCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCCTCCTGGTGGAACGCATCTACCACCTGCTGGAACGAGTCCCAGAACCAGAT
TGGAGTGGGAGAGCGCGAGATCGCGGACAACAAGTGGGCCTCCCTGTGGAACTGGTTCGACATCACCAACTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGG
CTGGTGGGCCTGCGGATCTGCATCGTGTTCGCCGTGCTGTCCATCGTGAACCGCGTGCGCCAGGGCTACTCCCCGTCCTTCCAGACCCGCCTGCCCGCC
TGCCCGCCCCCCGCGACCGCCCTGCCCTCCGGCGAGGGCGACCGGGCGGCGAGCATCGAGGAGGGCGGCGGAGCGCGCGGCCGGGGACCGGC
TTCCTGGCCCCTGATCTGGGACGACCTGCGCTCCCTGTGCCTGTTCTCCTACCACCGGCTGCGCGACCTGCTGCTGATCGTGACCCGCATCGT
GGAGCTGCTGGGCCGCCGCGGGGAGGTGCTGAAGTACTGGTGGAACCTGCTGCAGTACTGGTCCCAGGAGCTGAAGAACTCCGCGTGT
CCCTGCTGAACGCCACCGCCATCGCCGTGGCGCGAGGGCACCGACCGGGTTGATCGAGGTGGTGCAGCGCGTGTTCCGCGCCATCCTGCACATC
CCCCGCCGCATCCGCCAGGGCCTGGAGCGGCGGGCCCTGCTGTAA
```

Fig. 45A

2003 B.anc Env

MRVKGIRKNCQHLWRWGTMLLGMLMICSAAENLWVTVYYGVPVWKEATTLFCASDAKAYETEVHNVWATHACVPTDPNPQEVVLENVTENF
NMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLLNATNTNSTNMYRWRGEIKNCSENITTSIRDKMQKEYALFYKLDVVPIDNN
TSYRLINCNTSVITQACPKVSFEPIPIHYCTPAGFAILKCNDKKFENGTGPCKNVSTVQCTHGIRPVVSTQLLLNGSLAEEEVIRSENFTDN
AKTIIVQLNESVEINCTRPNNNTRKSIHIGPGRAFYATGEIIGDIRQAHCNLSRAKWNNTLKQVVTKLREQFDNKTIVFNPSSGGDPEIVMH
SFNCGGEFFYCNTTQLFNSTWNGTWNNTEGNITLPCRIKQIINMWQEVGKAMYAPPIRGQIRCSSNITGLLLTRDGGNNETEIFRPGGGDMR
DNWRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRAVGIGAMFLGFLGAAGSTMGAASMTLTVQARQLLSGIVQQQNNLLRAIEAQQHLLQL
TVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICTTVPWNASWSNKSLDEIWNNMTWMEWEREIDNYTGLIYTLIEESQNQQEKNEQEL
LELDKWASLWNWFDITNWLWYIKIFIMIVGGLVGLRIVFAVLSIVNRVRQGYSPLSFQTRLPAPRGPDRPEGIEEEGGERDRSGRLVNGF
LALIWDDLRSLCLFSYHRLRDLLLIVARIVELLGRRGWEALKYWWNLLQYWSQELKNSAVSLLNATAIAVAEGTDRVIEVVQRACRAILHIP
RRIRQGLERALLS

Fig. 46A

2003 CON_C Env

MRVRGILRNCQQWWIWGILGFWMLMICNVVGNLWVTVYYGVPVWKEAKTLFCASDAKAYEKEVHNVWATHACVPTDPNPQEIVLENVTENF
NMWKNDMVDQMHEDIISLWDQSLKPCVKLTPLCVTLNCTNATNATNTMGEIKNCSFNITTELRDKKQKVYALFYRLDIVPLNENNSYRLINC
NTSAITQACPKVSFDPIPIHYCAPAGYAILKCNNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTNNAKTIIVHL
NESVEIVCTRPNNNTRKSIRIGPGQTFYATGDIIGDIRQAHCNISEDKWNKTLQKVSKKLKEHFPNKTIKFEPSSGGDLEITTHSFNCRGEF
FYCNTSKLFNSTYNSTNSTITLPCRIKQIINMWQEVGRAMYAPPIAGNITCKSNITGLLLTRDGGKNNTETFRPGGGDMRDNWRSELYKYKV
VEIKPLGIAPTKAKRRVVEREKRAVGIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQTRV
LAIERYLKDQQLLGIWGCSGKLICTTAVPWNSSWSNKSQEDIWDNMTWMQWDREISNYTDTIYRLLEDSQNQQEKNEKDLLALDSWKNLWNW
FDITNWLWYIKIFIMIVGGLIGLRIIFAVLSIVNRVRQGYSPLSFQTLTPNPRGPDRLGRIEEEGGEQDRDRSIRLVSGFLALAWDDLRSLC
LESYHRLRDFILIAARAVELLGRSSLRGLQRGWEALKYLGSLVQYWGLELKKSAISLLDTIAIAVAEGTDRIIELIQRICRAIRNIPRRIRQ
GFEAALQS

Fig. 45B

2003 B.anc Env.seq.opt

ATGGCGGTGAAGGGCATCCGCAAGAACTGCCAGCACCTGTGGCGCTGGGGCACCATGCTGTGCTGGGGCATGCTGATGATCTGCTCCGCCGCCGA
GAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCACCACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGAGA
CCGAGGTGCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGGTGGTGCTGGAGAACGTGACCGAGAACTTC
AACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCC
CCTGTGCGTGACCCTGAACTGCACCGACCTGCTGAACGCCACCAACACCAACAACAGCTCCGAAGGGAGATCAAGAACT
GCTCCTTCAACATCACCACCTCCATCCGCGACAAGATGCAGAAGGAGTACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACAACAAC
ACCTCCTACCGCCTGATCAACTGCAACACCTCCGTGATCACCCAGGCCTGCCCCAAGGTGTCCTTCGAGCCCATCCCCATCCACTACTGCAC
CCCCGCCGGCTTCGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACCGGCCCCTGCCGGAGGAGGTGGTCACCGTGCACCCACG
GCATCCGCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGGTGGTGATCCGCTCCGAGAACTTCACCGACAAC
GCCAAGACCATCATCGTGCAGCTGAACGAGTCCGTGGAGATCAACTGCACCCGCCCGAACAACACCCGCAAGTCCATCCACATCGGCCCC
GGCCGCGCCTTCTACGCCACCGGCGATATCATCGGCGACATCCGCCAGGCCCACTGCAACCTGTCCCGGCGGGCGACCCCGAGATCGTGATGCAC
AGCAGGTGGTGACCAAGCTGCGCGAGCAGTTCGGCAACAAGACCATCGTCTTCTACTGCAACCCCTGTTCAACTCCAACACCGAGGGCAA
CATCACCCTGCCCTGCCGCATCAAGCAGATCATCAACATGTGGCAGGAGGTGGGCAAGGCCATGTACGCCCCCCCATCCGCGGCGGGCAGATCC
GCTGCTCCTCCAACATCACCGGCCTGCTGCTGACCCGCGATGGCGGCAACAACGAGACCGAGATCTTCCGCCCGGGCGGCGGCGGCGACATGCGC
GACAACTGGCGCTCCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCGGCGGCCAAGCGCGCCGTGGTGCA
GCGCGAGAAGCGCGCCGTGGGCATCGGGCCAGCTGCTGTTCCGGCATCGTGCCGGCGATCGTGCCGCGCCTCCATGACCCTGA
CCGGTGTGGGGCGGCCATCAAGCAGCTGCAGGCGCGGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGCGCGACCAGCAGCTGCTGGGCATCTGGGGCTGCTC
CGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACGCCTCCTGGTCCAACAAGTCCCTGGAGCAGATCTGGAACAACATGACCTGGATGG
AGTGGGAGCGCGAGATCGACAACTACACCGACCTGATCTACACCCTGCTGGAAGAGTCCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTG
CTGGAGCTGGACAAGTGGCCCTCCCTGTGGAACTGGTTCGACATCACCAACTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGG
CCTGGTGGGGGCCCCGCGGGCATCGTGTTCGCCGTGCTGTCCATCGTGAACCGCGTGCGCCAGGGCTACTCCCCCCTGTCCTTCCAGACCCGCCTGC
CCGCCCCCCCGCGACCGGCCCCCCGAGGGCATCGAGGAGGGCGGCGGGCGAGCGCGACCGCGACCGCTCCGGGCGTGCTGGTGAACGGCTTC
CTGGCCCTGATCTGGGACGACCTGCGCTCCCTGTGCCTGTTCTCCTACCACCGCCTGCGCGACCTGCTGCTGATCGTGACGCGTGCTGGTGGA
GCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGTGGAACCTGCTGCAGTACTGGGGCCAGGAGCTGAAGAACTCCGCCGTGTCCCC
CGCCCGCGCATCCTGCCGCGCCCCATCCTGCGCGCCCTGCAGACCTGCTGATCGTGGGCCGCGCCCATCGTGGGCCCGTGCACATCCCC
CGCCCGCCGATCCTGCCAGGGCCTGGCCGGAGGGCCTGGGGAGCGCGTGGGCCGACCGCGCCACCGCCCGCGCCCATCCTGCCGCGCGCCCTGCACATCCCC
CTGGCCCCTGATCTGGGACGACCTGCGCTCCCTGTGCCTGTTCTCCTACCACCGCCTGCGCGACCTGCTGCTGATCGTGGGCCGCGTGCTGGTGGA
GCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGTGGAACCTGCTGCAGTACTGGGGCCAGGAGCTGAAGAACTCCGCCGTGTCCCC
CGCCCGCGCCATCCTGCCAGGGCCTGGCCGGAGGGGCCTGGGGAGCGCGTGGGCCGACCGCGCCACCGCCCGCGCCCATCCTGCCGCGCGCCCTGCACATCCCC
CGCCCGCCGATCCTGCTGTAA

Fig. 46B

2003 CON_C Env.seq.opt

```
ATGCGCGTGCGCGGCATCCTGCCAGCAGTGGTGCAACTGCCACTGCTGTGGGGATCTGGGCGCATCCTGGGCTTCTGGATGCTGCTTCTGATGATCTGCAACGTGGTGGG
CAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAAGACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGAGA
AGGAGGTGCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCATCCTCCCTGTGGGACCAGTCCCTGAAGCCCTGAAGCTGACCCC
AACATGTGGGAAGAACGACATGGTGGACCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCC
CCTGTGTGCTGGACACCCTGAACTGCAACGCCACCAACGCCACCAACAACGCCACCAACAACGGCGAGATCAAGAACTGCTCCTTCAACATCACCACCGAGC
TGCGCGACAAGAAGCAGAAGGTGTACGCCCTGTTCTACCGCCTGGACATCGTGCCCCTGAACGAGAACAACTCCTACCGCCTGATCAACTGC
AACACCTCCGCCATCACCCAGGCCTGCCCCAAGGTGTCCTTCGACCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTACGCCATCCTGAA
GTGCAACAACAAGACCTTCAACGGCACCGGCCCCTGCCACAACGTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGTCCACCC
AGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGGAGATCATCATCCGCTCCGAGAACCTGACCGACAACGCCAAGACCATCATCGTGCACCTG
AACGAGTCCGTGGAGATCGTGTGCACCCGCCCCAACAACAACACCCGCAAGAGCATCCGCATCGGCCCAGGCCAGGCCTTCTACGCCACCCGG
CGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCTCCGAGGACAAGTGGAACAAGACCCTGCAGAGGGTGTCCAAGAAGCTGAAGG
AGCACTTCCCCAACAAGACCATCAAGTTCGAGCCCTCCTCCGGCGGCGACCTGGAGATCACCACCCACTCCTTCAACTGCCGCGGCGAGTTC
TTCTACTGCAACACCTCCAAGCTGTTCAACTCCACCTACACGCCCCCCATGCCCGGCGCAACATCACCCTGCCCTGCCGCATCAAGCAGATCATCAA
CATGTGGCAGGAGGTGGGCCGCGCCATGTACGCCCCGCCCCCCGCCGGCGGCGGCGCCTCAAGAGGCCGCTGCCTGCTGACCC
GCGACGGCGGCAAGAACAACACCGAGACCTTCCGCCCCGGCGGCGGCGGCGGCAAGGCCAAGCCCGCGCTGGTGGAGGCGGCCGTGGCGCGCCCGTT
GTGGGAGATCAAGCCCTGGGCCGCCGGCGGCGGCGGCCTCCACCATGGGCGCGCCCAGGCCCAGCAGCAGCACCTGCTGTCCGGCATCGTGC
CCTGGGCTTCCTGGGCGCGCCCGGCTCCGGCCCAGCCATCGAGGGCCGCGCCAGGCCCAGCAGCACCTGTGGGGCATCAAGCAGCTGCAGGCCCGTG
AGCAGCAGCAGCAGCCGCTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGGCAAGCTGATCTGCACCACCGCCGTGCCGTGGAA
CTGGCCCATCGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCGTCTGGGGCATCTGGGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAA
CTCCTCCTGGTCCAACAAGTCCCAGGACGACATCTGGGACAACATGACCTGGATGCAGTGGGACCGCGAGATCTCCAACTACACCGACACCA
TCTACCGCCTGCTGGAGGACTCCCAGAACCAGCAGGAGAAGAACGAGAAGGACCTGCTGGCCCTGGACTCCTGGAAGAACCTGTGGAACTGG
TTCGACATCACCAACTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGATCGGCCTGCGCATCGTGTTCGCCGTGCTGTC
CATCGTGAACCGCGTGCGCCAGGGCTACTCCCCCCTGTCCTTCCAGACCCTGTTCCCCGCCCCGCGCGGCCCTGGACCGCCTGGGCCGCATCGAG
AGGAGGGGCGGCGAGCAGGACGGCGGCGAGCGCGACCGCGAGCGCCTCCCCCGCGCGCTCCCCCCATCCCCCTGGGCGCGCCCTGGAGGCTGGACACCAGCG
CTGTTCTCCTACCACCGGCTGCGCGACTTCATCCTGCTGGTGGTGGCCCGCATCGTGGAGCTGCTGGGCCGCCGCGGCTGGGAGGCCCTGAAGTACTGGTGGAAC
CGGCTGGGCTACTGGGGCCGCGAGCTGAAGAACTCCGCCGTCTCCCTGCTGGACACCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCATCCTGGAGCTGATCCAGCGCATCTGCCGCGCCATCCTGAACATCCCGCGCCGCATCCGCCAGGGCGCCGAGCGCGCCCTGCTGTAA
GGCTTCGAGGGCGCCCCTGCAGTAA
```

Fig. 47A

2003 C.anc Env

```
MRVMGILRNCQQWWIWGILGFWMLMICNVVGNLWVTVYYGVPVWKEAKTTLFCASDAKAYEREVHNVWATHACVPTDPNPQEMVLENVTENF
NMWKNDMVDQMHEDIISLWDQSLKPCVKLTPLCVTLNCTNATNATNTMGEMKNCSFNITTELRDKKQKVYALFYRLDIVPLNDNNSYRLINC
NTSAITQACPKVSFDPIPIHYCAPAGYAILKCNNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAEEIIIRSENLTDNAKTIIVHL
NESVEIVCTRPNNNTRKSIRIGPGQTFYATGDIIGDIRQAHCNISEEKWNKTLQRVGEKLKEHFPNKTIKFAPSSGGDLEITTHSFNCRGEF
FYCNTSRLFNSTYNSKNSTITLPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGKNNTETFRPGGGDMRDNWRSELYKYKV
VEIKPLGIAPTEAKRRVVEREKRAVGIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQTRV
LAIERYLKDQQLLGIWGCSGKLICTTAVPWNSSWSNKSQEEIWDNMTWMQWDREISNYTDTIYRLLEDSQNQQEKNEQDLLALDSWENLWNW
FDITNWLWYIKIFIMIVGGLIGLRIIFAVLSIVNRVRQGYSPLSFQTLTPNPRGPDRLGRIEEEGGEQDRDRSIRLVSGFLALAWDDLRSLC
LFSYHRLRDFILIAARAVELLGRSSLRGLQRGWEALKYLGSLVQYWGLELKKSAISLLDTIAIAVAEGTDRIIELIQRICRAIRNIPRRIRQ
GFEAALL$
```

Fig. 48A

2003 CON_D Env

```
MRVRGIQRNYQHLWRWGIMLLGMLMICSVAENLWVTVYYGVPVWKEATTLFCASDAKAYDTEAHNIWATHACVPTDPNPQEIELENVTENF
NMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVKRNNTSNDTNEGEMKNCSFNITTEIRDKKQKVYALFYKLDVVPIDDNNSNT
SYRLINCNTSAITQACPKVTFEPIPIHYCAPAGFAILKCKDKKFNGTGPCKNVSTVQCTHGIRPVVSTQLLLNGSLAEEEIIIRSENLTNNA
KIIIVQLNESVTINCTRPYNNTRQRTPIGPGQALYTTRIKGDIRQAHCNISRAEWNKTLQQVAKKLGDLLNKTIIFKPSSGGDPEITTHSF
NCGGEFFYCNTSRLFNSTWNNTKWNSTGKITLPCRIKQIINMWQGVGKAMYAPPIEGLIKCSSNITGLLLTRDGGANNSHNETFRPGGGDMR
DNWRSELYKYKVVKIEPLGVAPTRAKRRVVEREKRAIGLGAMFLGFLGAAGSTMGAASMTLTVQARQLLSGIVQQQNNLLRAIEAQQHLLQL
TVWGIKQLQARILAVERYLKDQQLLGIWGCSGKHICTTTVPWNSSWSNKSLDEIWNNMTWMEWEREIDNYTGLIYSLIEESQNQQEKNEQEL
LELDKWASLWNWFSITQWLWYIKIFIMIVGGLIGLRIVFAVLSIVNRVRQGYSPLSFQTLLPAPRGPDRPEGIEEEGEQGRGRSIRLVNGF
SALIWDDLRNLCLFSYHRLRDLILIAARIVELLGRRGWEALKYLWNLLQYWIQELKNSAISLFDTTAIAVAEGTDRVIEIVQRACRAILNIP
TRIRQGLERALL$
```

Fig. 47B

2003 C.anc Env.seq.opt

ATGCGCGGTGATGGGCATCCTGCGCAACTGCCAGCAGTGGTGGATCTGGGGCATCCTGGGCTTCTGGATGCTGGTTCCTGGCGTGCTGATGATCTGCAACGTGGTGGG
CAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAAGACCACCCTGTTCTGCGCCTCCAGGAGATGGTGCTGGAGAACGTGACCGAGAACTTC
GCGAGGTGCACAACGTGTGGGCCACCCACGCGCCTGCGTGCCACCAGATGGACCAGATCACGAGGACCATCATCTCCCTGTGGGACCAGTCCCTGAAGCCCCTGAAGCTGACCCC
AACATGTGGCGGTGAACGACACATGGTGCACCAGATGACGGTGCACCAACAGCCAACAACGCCAACAACTGCTCCTTCAACATCACCACCGAGC
CCTGTGTGCCTGACCCTGACCCTGAACTGCGTCACCAACAACTGTCTCCTTCAACATCGTGCCCCTGACACTGCCTGATCAACTGC
TGCCGCGACAAGAAGCAGAAGGTGTACGCCCTGTTCTACCGCCTGGACATCGTGCCCCTGAACGACAACAACTACTGCGCCCCGGCTACGCCATCCTGAA
AACACCTCCGCCATCAACAAGAGACCTTCAACGGCACCGACCGAGTGCTGCGCCGGCTACGCCATCCTGAAGCCCGGCTACGCCATCCTGAA
GTGCAACAACAAGACCTTCAACGGCACCTCCACGGCACCCGCCCCGAGGAGGAGAATCATCATCTCCGTGCAGTGTCCAGCTGAA
AGCTGCTGCTGAACGGCTCCCTGGCCTGCGTGTGCACCCGCCATCGTGCGCCCACTGCAACATCTCCGAGGAGGAGCTGGCGCGGAAGCTGAAGG
CGACACTTCCCAACAAGACCATCAAGTTCGCCCCCTCGCCCCTGTTCAACTCCAACACCTCCACCACCGAGATCCACAGCAGATCAA
TTCTACTGCAACACACCTCCCGGCCTGTTCAACTCCAACATCGTGCCCCATCAAGCAGATCAA
CATGTGTGGCAGGGCGTGGGCCGCGCGCCCATGTACGCCCCCCCGGCCTGGGGCGCGGAAGCTGCTGACCCC
GCGCGGCGGGCCAAGAACAACAACACCGAGACCTTCCGCCCCGGCGGCGGGCGACTACAAGGTG
GTGGAGATCAAGCCCCTGGGGCATCGCCCGCTCCACCACCGAGAAGCGCCGCCGCAGCTGCGTGTCCCGGCATCGCGTGCC
CCTGGGCTTCCTGGGCGCGCCTGCGGCCCCGCCATGCGAGGCCCAGCAGCACATGTCTGGGCGTGTGGGCTGATCTGCAA
AGCAGCAGTCCCATCGCCCTGCTGCTCAGAAGGACCAGCAGCTGTTGGGGACAACATGACCGAGACCGCCCTGACCGCCCTGGAACTGG
CTGGCCCATCGAGCCGCTGCGTGGTCCAACAAGTCCCAGGAGGACGAGAGATCTTCATCATGACCGTGCTGCCCGTGTCC
CTCCCTCCTGGTCCGCCTGGTGGAGGACTCCCAGAACCAGCAGGAAGGAGAAGCCGTGCCTGGAACTGG
TTCGACATCACCAACTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGATCGGCCTGCGTGTC
CATCGTGAACGGCGTGCGCCAGGGCTACTCCCCCCTGTCCTTCCAGACCCTGCTGCCCGCGCCGAGCCTGCGCGCCATCG
AGGAGGGCGGCGGCGAGCAGGAGCGCGGCCGCTCCATCCGCCTGGTGGGCGGCGCCGCGCTCCCCTGTGC
CTGTTCTTCCTGTACCACGCGCCTGCGCGACTTCATCCTGATCGCCGCGCTGCGCGCCGGCGCGCGCAGCG
CGGCTGGGAGGCCCTGAAGTACTGGTGGGCTCCCTGGGTCAGTACTGGGGCCCTGCAGTACTGGGGCCCTGGAGCTGGGACACCATG
CCATCGCCGTGGCCGCTGCAGGCTACCGGGCCATCATCTGCGCGCGCCACCATCCCGCGGCCATCGCCGCGCGCCAG
GGCTTCGAGGGCCGCCCTGCTGTAA

Fig. 48B

2003 CON_D Env.seq.opt

ATGCGCGGTGCGCGGCCATCCAGCGCCAACTACCAGCACCTGTGGGCTGCTGGGCATCATGCTGCTGGGCATCTGCTCCGTGGCCGA
GAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCACCACCACCCTGTTCTGCGCCTCCAAGTCCTACAAGA
CCGAGGCCCACAACATCTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCCAGGAGATCAGCCTCCCTGTGGGACACCATCTCCCTGTGGGACCAGTCCCTGAAGCCCTGAAGCTGACCCC
CCTGTGCGTGACCCTGAACTGCACCGATGCCAACTGGCAGAGTGAAGGCGAACAACACCTCCAACGACACCAACAGAGGGCGAGATGAAGAACTGCTCCTTCAACA
TCACCACCGAGATCCGCGACAAGAAGCAGGTGCACGCCCTGTTCTACAAGCTGGACGTGGTGCCCATCGACGACGACAACTCCAACACC
TCCTACCGCCTTCGCCATCCTGATCAACTGCAACACCTCCGCCACCAACCTCCAGGCCCCCCAAGGTGACCTTCGAGCCCATCCCCATCCACTACTGCGCCCC
CGCCGGCTTCGCCATCCTGAAGTGCAAGGACAAGAAGTTCAACGGCACCCTCCCTGGCCGAGGAGGAGATCATCATCCGCAGCGAGAACGTGTCCGCTCCGAGAACCTGACCAACAACGCC
AAGATCATCATCGTGCAGCTGAACGAGTCCGTGACCATCAACCGCCCCACTGCAACATCTCCGCGGCGGACAACACTCTCCGCGGCGGACAACCACCCTGCAGGAGG
CCAGGCCCTGTACACCACCGGCGACCTGCTGAACAAGAACATCTTCAAGCCCTCCCACCTGGAACAACACCAAGTGGAACTCCACCGGCAAGATCACCAGTGCT
TGGCCAAGAAGCTGGGCGAGTTCTTCTACTGCAACAGCACCAACCTGTTCAACTCAACAACATGTGGCAGGGGCCGCGGCGCGCCAACAACTCCCCAACAACTGAGCCCTGGGCGGGGGGACATGCGC
AACTGCGGCGGCGAGTTCTTCTACTGCAACAAGACCAACATCTCCGCCTGTTCAACTCTGTTCAACTCCGGCAGGGGCTGGCAGGGCGCGGCGCGCCAACAACTCCCCAACAACTGAGCCCCATCGAGGTGGCCAATCGAGGCTGATCAAGTGCT
CCTGCCCCCTGGGCCGAGTTCTTCTACTGCAACAGCACCAACATCAACAACATGTGGCAGGGCGCGGCGCGCCAACAACTCCCCAACAACTGAGCCCCATGCCCCGGCCGGCCAAGCGGCGCGGCGTGGA
GACAACTGGCCTCCGAGCTGTACAAGTACAAGGTGGTGAACAACTCGGGCTTCCTGGCGCGCCATGTTCCTGGCGCCATGTTCCTGCAGCAGAACAACCTGCTCGCGCGCCAGCAGAGCGCGGCGCGCCAGCAGGCGCGGTGGA
GCGCCGAGAAGCGGCCCGCGCCATCGGCCTGGGCGCGCTGTTCCTGGGCCATGTTCCGGCGCATCGTCGTGCAGCAGAACAACCTGGGCGCGCGCCAACAACTGCTGCTGGGCTGCTGCTGGGCTGCTGCTGA
CCGGTGCAGGAAGCGCCCCGCGCGCCATCAGCGACCTGCAAGCAGCTGCGCCCTGGGCGCCATCTCCTGGAACTCCTGGGACACGGGAACTGGAACATGACCTGCTGGGGCTGCTC
AGTGGGAGCGCGGAGATCGACAAGTGGGCCTCCCTGTGGACCTGGTTCTCCATCACCAACCTGGTGCTGGTTCTCCATCACCAACCTGGTGCTGTTCCGCCCTGGGCGCTGGGCGCTG
CTGGGAGCTGGACAAGTGGGCCTCCCTGTGGACCTGGTTCTCCATCACCAACCTGGTGCTGGTTCTCCATCACCAACCTGGTGCTGTTCCAGACCCTGAACCGGCTTC
CCTGATCGGCCTGCGCATCGTGTTCGCCGTGCTGAGCATCGTGAACCGCGTGAGGGGCATCCGCCAACCTGTGCCGCGCCTGGGCGCCTGATCCTGGAACGGCTTC
CCGCCCCCTGGGACCGCCCCTCGTTCGTGCAGCTGCCGGCCAACCTGTCCTGATCGGGGCGGCCTGATCCTGGGGCGCGGCGCGCCGCGCGGCCATCCGCCATCCCC
TCCGCCCTGATCTGGGACGACCTGCGCAACCTGTGCCTGTTCTCCTACCACCGCCTGCGCGACCTGCTGCTGATCGTGACCCGGCGCCGCGCGCCAGCGGCGCCAGCAGTACCTGCCGGGA
GCTGCTGGGCAACCTGCGCCCATCCGCCATCCCGGAGGGCTGGGAGAGGGCCATCGCCCGGGCGCGGCGGCAACCTGTGCCTGTTCTCCTACCACCGCCTGCGCGACCTGCTGCTGATCGTGACCCGGCGCCGCGCGCCAGCGGCGCCAGCAGTACCTGGGAACTCGTGGA
TGTTCGACACCAACCGGCCATCCGCCATCCCGGAGGGCTGGGAGAGGGCCACCGCCCGGGCGCGGCGGCAACCTGTGCCTGTTCTCCTACCACCGCCTGCGCGACCTGCTGCTGATCGTGACCCGGCGCCGCGCGCCAGCGGCGCCAGCAGTACCTGAGATCGTGCAGCGCGCCATCCGCGCCATCCCC
ACCCGCGCATCCGCCAGGGCCTGGGAGGGCGCGGCCCCTGCTGTAA

SUBSTITUTE SHEET (RULE 26)

Fig. 49A

2003 CON_F1 Env

MRVRGMQRNWQHLGKWGLLFLGILIICNAAENLWVTVYYGVPVWKEATTLFCASDAKSYEKEVHNVWATHACVPTDPNPQEVVLENVTENF
DMWKNNMVEQMHTDIISLMDQSLKPCVKLTPLCVTLNCTDVNATNNDTNDNKTGAIQNCSFNMTTEVRDKKLKVHALFYKLDIVPISNNNSK
YRLINCNTSTITQACPKVSWDPIPIHYCAPAGYAILKCNDKRFNGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEDIIIRSQNISDNAK
TIIVHLNESVQINCTRPNNNTRKSIHLGPGQAFYATGEIIGDIRKAHCNISGTQWNKTLEQVKAKLKSHFPNKTIKFNSSSGGDLEITMHSF
NCRGEFFYCNTSGLFNDTGSNGTITLPCRIKQIVNMWQEVGRAMYAAPIAGNITCNSNITGLLLTRDGGQNNTETFRPGGGNMKDNWRSELY
KYKVVEIEPLGVAPTKAKRQVVKRERRAVGIGAVELGFLGAAGSTMGAASITLTVQARQLLSGIVQQQNLLRAIEAQQHLLQLTVWGIKQL
QARVLAVERYLKDQQLLGIWGCSGKLICTTNVPWNSSWSNKSQDEIWNNMTWMEWEKEISNYSNIIYRLIEESQNQQEKNEQELLALDKWAS
LWNWFDISNWLWYIKIFIMIVGGLIGLRIVFAVLSIVNRVRKGYSPLSLQTLIPSPREPDRPEGIEEGGEQGKDRSVRLVNGFLALVWDDL
RNLCLFSYRHLRDFLLIAARIVDRGLRRGWEALKYLGNLTQYWSQELKNSAISLLNTTAIVVAEGTDRVIEALQRAGRAVLNIPRRIRQGLE
RALLS

Fig. 50A

2003 CON_F2 Env

MRVREMQRNWQHLGKWGLLFLGILIICNAADNLWVTVYYGVPVWKEATTLFCASDAKAYEREVHNVWATYACVPTDPSPQELVLGNVTENF
NMWKNNMVDQMHEDIISLMDQSLKPCVKLTPLCVTLNCTDVNVTINTTNVTLGEIKNCSFNITTEIKDKKKEYALFYRLDVVPINNSIVYR
LISCNTSTVTQACPKVSFEPIPIHYCAPAGFAILKCNDKKFNGTGLCRNVSTVQCTHGIRPVVSTQLLLNGSLAEEDIIIRSENISDNTKTI
IVQFNRSVEINCTRPNNNTRKSIRIGPGRAFYATGDIIGDIRKAYCNINRTLWNETLKKVAEEFKNHFNITVTFNPSSGGDLEITTHSFENCR
GEFFYCNTSDLFNNTEVNNTKTITLPCRIRQFVNMWQRVGRAMYAPPIAGQIQCNSNITGLLLTRDGGKNGSETLRPGGGDMRDNWRSELYK
YKVVKIEPLGVAPTKAKRQVVQREKRAVGIGAVLLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLKAIEAQQHLLQLTVWGIKQLQ
ARILAVERYLKDQQLLGIWGCSGKLICTTNVPWNSSWSNKSQDEIWDNMTWMQWEKEISNYTDTIYRLIEDAQNQQEKNEQDLLALDKWDNL
WSWFTITNWLWYIKIFIMIVGGLIGLRIVFAVLSVVNRVRQGYSPLSLQTLIPNPRGPERPGGIEEGGEQDRDRSIRLVSGFLALAWDDLR
SLCLFSYRHLRDFLLIAARTVDMGLKRGWEALKYLWNLPQYWGQELKNSAISLLDTTAIAVAEGTDRIIEVLQRAGRAVLHIPRRIRQGFER
ALLS

Fig. 49B

2003 CON F1 Env.seq.opt
ATGGCGGTGCGCGGCCATGCAGCGGCAACTGGCAGCACCTGGGCGGCCTGCTGTTCCTGGGCATCATCCTGATCATCCTGCAACGCCGCCGA
GAACCTGTGGGTGACCGTGTACTACGGCGTGTGGGCCACACGCCTGCCACCCAACCCCCAGGAGGTGCTGGAGAACGTGACCGAGAACTTC
AGGAGGTGCACAACGTGCACAACATGGTGGAGCAGATGCACACCGACATCATCTCCCTGTGGGACCAGTCCCTGAAGCCCATCATCCTGCA
GACATGTGCTGACGTGAACTGACCGACGTGAACGCCACCAACCAACAAGACCGGCCGCCATCTCCAACAACAACTGCTCCTTCA
CCTGTGCTGACCGAGGTGCGCGACAAGAAGCTGAAGGTGCACGCCCTGTTCTACAAGTCTGGGACATCGTCCAACAACTGCTCCTTCA
ACATGACCACCGAGGTGCGCGACAAGAAGCTGAAGGTGCACGCCCTGTTCTACAAGTGTCCTGGGACATCGTCCAACAACTGCTCCTTCA
TACCGCCTGATCAACTGCAACTGCCATCCTGAAGTGCAACGGCTTCAACGCAAGGCTTCAAGCCCCTGGCCAGCTCCCTGAACGGCTCCCTGGCCGCA
CGGCTACGCGTCTGTCCACCCAGCTGCTGAACGGCTCCCAGCTGCTGAACGGCTCCCTGGCCGCGAGGAGGACAATCATCCTCCGGCGCGCGCAAG
AGCCTGGTGGTGCTCACCCAGCTGCTGAACGGCTCCCTGGCCGAGCTCCCAGAACATCTCGCACCTGGGCCCCGGCCAAG
ACCATCATCGTGCACCTGAACGAGTCCGTGCAGATCAACTGCACCCGGCCCAACAACAACAACCTCCTCCGGCGCGCGCAAG
GGCCTTCTACGCCGAGATCATCGGCGACATCCCCAACAAGCTGAACTGCACCCGGCCTCAACTCCAACATCCGCCGCATCAA
TGAAGGCCAAGCTGAAGTCCCACTTCCCCAACAAGACCATCAAGTTCAACAACCGGCCCAACAACCTGCCCAACATCACCGGCC
AACTGCCGCGGCGAGTTCTTCTACTGCAACACCACCTCCGGCCTGTTCAACGATACGGCCCCATGAGCGCCCTCAACATCACCGGCC
GCAGATCGTGAACATGTGGCAGGAGGTGGGCAAGGCCATGTACGCCCCCCCCGGCCGTGGCCGCGCCCAGCTGCTGT
TGCTGCTGACCCCGACGGCGCCAGAACAACACCGAGACCTTCCGCCCTGGCCGCCCTGCTGCGGCCCGCGCCAGCTGCTGT
AAGTACAAGGTGGTGACCCGCGAGCATGTGGCAGGAGGTGGGCAAGGCCATGTACGCCCCCCCCGGCCGTGGGGCAT
CGGCGGCCGTGTTCCTGGGCGTTCCTGGGCGGCCTTCCTGGGCGGCCGTGTACCTGCTGCTGCTGGGGCTGTGT
CCGGCATCGTGCAGCAGATCAACAACTGGCGCGGCCAGGACCAGAGCCAGTGCTGCTGGGGGCCTGTGGGGCATCAAGCAGCTG
CAGGCCCGCGTGCTGGCCGTTGCTGGCGCTACCTGAAGGACCAGACCAGAGCCAGTGCTGCTGGGGCTGTGTGGGGCAGCTGTGGGGCAGCTG
CGTGCCCCTGAACTCCTCCTGGTCCAACAAGTCCCTGATCGAGGAGATCGAGGAACAACAGCAGGAGAAGAAGGAGATCTCCAACT
ACTCCAACATCATCTACGGCCTGGTTCGAACCGCGTGGGCCAAGGGCTACTCCCCCCGCTGCCCCTGATCCCCCGCGAGCCCGACCGCC
CTGTGGGAACTGGTTCGACATCTCCAACTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGATCGGCCTGCGCATCGTGTT
CGCCGTGCTGTCTGTCATCGTGAACCGCGTGCGCCAAGGGCTACTCCCCCTGATCCCCCGCGAGCCCGACCGCC
CCGAGGGCATCGAGGAGGGCGGCGAGCGGCGACCGCCATCGTGAGCCCTGGTGTGGGGACCTG
GCAACCTGTGCCTGTTCTCCTACCGCCGACTTCATCCTGACTGGTTCCAGCCCCTGCCCCTGGGGCCGCCCATCGTGGGA
GGCCCTGAAGTACCTGGGCAACCTGACCCAGTACTGGTCGCTGGAGCTACTGGTTCCAGCCCTGGCACCACCGCCATCGTGGGA
TGGCCGAGGGCACCGACCGCGTGATCGAGGCCATTGTGGAGCTGAACCCTGCTGAACACCCCCGCGCATCCGCGGCCAGGGCCTGGAG
CGCGCCCCTGCTGTAA

Fig. 50B

2003 CON_F2_Env.seq.opt

```
ATGCGCGTGCGCGAGATGCAGCGCAACTGGCAGCACCTGGGCAAGTGGGGCCTGCTGCTGTTCCTGGCCATCATCCTGATCATCTGCAACGCCGCCGA
CAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCACCACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGAGC
GCGAGGTGCACAACGTGTGGGCCACCTACGCCTGCGTGCCCACCGACCCCTCCCCCCAGGAGCTGGTGCTGGGCAACGTGACCGAGAACTTC
AACATGTGGAAGAACAACATGGTGGACCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCC
CCTGTGCCTGACCCTGACCTGCACCGACCTGAACGTGACCAACACCACCAACGCCTGGGCGAGATCAACAACTGCTCCTTCAACA
TCACCACCGAGATCAAGGACAAGAAGCAGAAGGAGTACGCCCTGTTCTACCGCCTGGTGCCCATCAACAACTGCTCCATCGTGTACCGC
CTGATCTCCTGCAACACCTCCACCGTGACCCAGGCCTGCCCCAAGGTGTCCTTCGAGCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTT
CGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACCGGCCTGTGCCGCAACGTGTCCACCGTGCAGTGCACCCACGGCATCCGCCCCG
TGGTGTCCACCCAGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGGACATCATCATCCGCTCCGAGAACATCTCCGACAACACCCCGCGCCATC
ATCGTGCAGTTCAACCGCCGCGACATCATCGGCGACATCCGCCAGGCCTACTGCAACATCTCCGGGCGACAACATCAACACCCTGGAGATCTCCGCCA
CTACGCCACCGGCGACATCATCGGCGACACCTTCAACCCCTCCGGCGGCGACCCTGAACGACACCCCAGCGTGAACGAGATCCCCGCCATCGTGGCCG
AGGAGTTCAAGAACCACTTCAACCGCGTGAACCTCCGACCGTGAACACCACCCTGGAGATCAACACCGCCTCCTCAACTGCCCA
GGCGAGTTCTTCTACTGCAACACCTCCGACCTGTTCAACAACACCAGAGTGAACACCACCCTGGCCGGCGCCATCCGCGCCATCCGCCA
GTTCGTGAACATGTGGCAGGCCGTGGGCAAGGCCATGTACGCGCCCCCATCGGCGGGCCTCCGAGAACGGCTCCGAGCTGCTGTACAAG
TGCTGACCCGCGACGGCGGCAAGAACGGCTCCGAGAACGGCTCCCTGGCCGTGGCCCCCCTGGGGGCCGGCCTCCATCTGGGCATCGG
TACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCTCCCGGCGCCCAGAGGCCCAGGTGGTGGTGACCCTGCCAGCTGGGCATGGG
CGCCGTGTCCGGTGGTGAAACCGGCGTGCCCAGGGCGCCGAGAGCGCCCTCATCACCCTGCCAGCTGCTGACCGTGCTGGGCGCATCGG
GCATCGTGCAGCAGCAGCAGTCCAACCTGCTGCGGCCCATCGAGGCCCAGCAGCACCGTGTGGGCATCAAGCAGCTGCAG
GCCCGCCATCCTGGCCGTGGAGCGCTACCTGAAGGCGCTACCTGAAGGACCAGCAGCTGCTGGGCATCTGGGGCAGCTGGGACGTCAACAACGT
GCCCTGGAACTCCTCCTGGTCCAACAAGTCCCAGGAGAGAGCGCCCAGAACCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGACAAGTGGGACAACCTG
TGGTCCTGGTTCAACATCACCAACTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGGTCGGCCTGCGTGTTCGC
CGTGCTGTCCGTGGTGAACCGCGTGCGCAAGGGCTACTCCCCGCTGTCCTTCCAGACCCTGGTTCCCCGCACCGAGGCCCCCGAGCGCCCCG
GCGGCATCGAGGAGGAGGGCGGCGAGCAGGACCGCGACCGCTCCATCCGCCTGGTGAAGGGCTCCTGGACCGGCCTGGGGAGGC
TCCCTGTGCCTGTTCTCCTACCACCGCCTGCGGCGACACTTCATCCTGATCGTGGCCGCCAGGCTGAAGAACTCCGCCCACCGCCATCGCGCGC
CCTGAAGTACCTGTGGAACCTGCTGCAGTACTGGGGCCAGGAGCTGAAGAGCCTGTGCTGAAGTACCTGGGCAACCGGCTGGGAGGC
CCGAGGGCACCGACGCGATCATCGAGGTGCTGCAGAGAGCCTGGGGCCGCCGCAATCGAGGCATCCCCCGCCGCAATCCGGCCTTCGAGCGC
GCCCTGCTGTAA
```

Fig. 51A

2003 CON_G Env

MRVKGIQRNWQHLWKWGTLILGLVIICSASNNLWVTVYYGVPVWEDADTTLFCASDAKAYSTERHNVWATHACVPTDPNPQEITLENVTENF
NMWKNNMVEQMHEDIISLWDESLKPCVKLTPLCVTLNCTDVNVTNNNTNNTKKEIKNCSFNITTEIRDKKKEYALFYRLDVVPINDNGNSS
IYRLINCNVSTIKQACPKVTFDPIPIHYCAPAGFAILKCRDKKFENGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIRSENITDNT
KVIIVQLNETIEINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSRTKWNEMLQKVKAQLKKIFNKSITFNSSSGGDLEITTHSF
NCRGEFFYCNTSGLFNNSLLNSTNSTITLPCKIKQIVRMWQRVGQAMYAPPIAGNITCRSNITGLLLTRDGGNNNTETFRPGGGDMRDNWRS
ELYKYKVKIKPLGVAPTRARRRVVEREKRAVGLGAVLLGFLGAAGSTMGAASITLTVQVRQLLSGIVQQQSNLLRAIEAQQHLLQLTVWGI
KQLQARVLAVERYLKDQQLLGIWGCSGKLICTTNVPWNTSWSNKSYNEIWDNMTWIEWEREISNYTQQIYSLIEESQNQQEKNEQDLLALDK
WASLWNWFDITKWLWYIKIFIMIVGGLIGLRIVFAVLSIVNRVRQGYSPLSFQTLTHHQREPDRPERIEEGGGEQDKDRSIRLVSGFLALAW
DDLRSLCLFSYHRLRDFILIAARTVELLGRSSLKGLRLGWEGLKYLWNLLLYWGQELKNSAINLLDTIAIAVANWTDRVIEVAQRACRAILN
IPRRIRQGLERALLS

Fig. 52A

2003 CON_H Env

TRVMETQRNYPSLWRWGTLILGMLLICSAAGNLWVTVYYGVPVWKEAKTTLFCASDAKAYETEKHNVWATHACVPTDPNPQEMVLENVTENF
NMWENDMVEQMHTDIISLWDQSLKPCVKLTPLCVTLDCSNVNTTNATNSRFNMQEELTNCSFNVTTVIRDKQQKVHALFYRLDVVPIDDNNS
YQYRLINCNTSVITQACPKVSFEPIPIHYCAPAGFAILKCNNKTFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAEEQVIIRSKNISDN
TKNIIVQLNKPVEITCTRPNNNTRKSIHLGPGQAFYATGDIIGDIRQAHCNISGKKWNKTLHQVVTQLGKYFDNRTIIFKPHSGGDMEVTTH
SFNCRGEFFYCNTSGLFNSSWTNSTNDTKNIITLPCRIKQIVNMWQRVGQAMYAPPIKGNITCVSNITGLILTFDEGNNTVTFRPGGGDMRD
NWRSELYKYKVVKIEPLGVAPTEARRRVVEREKRAVGMGAFFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLRAIQAQQHMLQLT
VWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTNVPWNSSWSNKSLDEIWDNMTWMEWDKQINNYTEEIYRLLEVSQTQQEKNEQDLL
ALDKWASLWNWFSITNWLWYIKIFIMIVGGLIGLRIIFAVLSIVNRVRQGYSPLSFQTLIPNPRGPDRPEGIEEGGEQDRDRSVRLVNGFL
PLVWDDLRSLCLFSYRLLRDLLLIVVRTVELLGRRGREALKYLWNLLQYWGQELKNSAINLNTTAIAVAEGTDRIIEIVQRAWRAILHIPR
RIRQGFERTLLS

SUBSTITUTE SHEET (RULE 26)

Fig. 51B

2003 CON G Env.seq.opt

```
ATGGCGGTGAAGGGCATCCAGCGCCAACTGGCCAGCACCTGTGGAAGTGGGGCACCCGTGGGACCCTGATCCTGGGCCTGGTGATCATCTGCTCCGCCTCCAA
CAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGGAGGACGCCGACACCACCCTGTTCTGCGCCGAGAACGTGACCGAGAACTTCCA
CCGAGCGCCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACATCATCTCCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCC
AACATGTGGCAGAACAACATGGTGGAGCAGATGCACCGAGATCTGACCAACAACAACAACACCAAGAAGGAGATCAAGAACTGCTCCTTCAACA
CCTGTGCCGTGACCCTGAACTGCACCGACGTGACCAACACCAAGACCCCTGTTCTACCGCCCCAAGGTGGTGCCCATCAACGACACCAACGCAACTCCTCC
TCACCACCGAGATCCGCGACAAGAAGAAGAAGAGTACGCCCTGTTCTACCGCCTGGACGTGGTGCCCATCAACGACACCAACGCCAACTCCTCC
ATCTACCGCCTGATCAACTGCAACGTGTCCCGCGACATCAAGCAGGCCTGGTTCGACCTGCTGGTGGTGCCCATCTACCTGCGCCCC
CGCCGGCCTTCGCACCGTGTCCCAGCTGCTGAACCGGCTCCCTGGCCGAGGAGGAGATCATCATCCCGAGAACATCACCGACAACACC
TCAAGCCCGTGGTGTCCCAGCTGCTGAACGAGATCAACACCGCCACCGCCCCGCCACCCGGCCCCCGG
AAGGTGATCATCGTGCAGCTGAACGAGACCATCGGCGCACATCAACTCCTGGAACTGGAAGCGAGATGCTGCAGA
CCAGGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCAGCTCCTTCAACAAGTCCATCGTCCCTGGAGATCACCACCACCTGCCCTG
AGGTGAAGGCCCAGCTGAAGAAGATCTTCAACAGCTGCTTCAACAGTCCATCCTGGAGATCACCACCATCACCTGCCCTG
AACTGCCGCGGCGAGTTCTTCTACTGCAACACCTCCGGCCTGTTCAACAACTCCCTGCGCCATCACCTGCGGCTCCAACA
CAAGATCAAGCAGATCGTGCCGCATGTGGCAGGCCGTGGCAGCCGGCATGTGGCAGGCCGTGGCAGCCGGCCAGCGTGCCC
TCAGCGCCTGCTGCTGACCCGGCGGCAACAACACCACCGAGACCTTCCGCCCCCGGGAGGCGAGGAAGCGGC
GAGCTGTACAAGTACAAGATCGTGAAGATCGAGCTGTCTGCGGCTTCCTGGGCGCCGCCGCCGCTCCATCGGCGGTGTGGTGGGGCATC
CGTGGGCCTGGGCCTGCTGCTGCGCCTGCGCGCTGCTGCTGCCCATCGGCCTGCTGCGCGGCCAGCGTGTGGGGCATC
AGCAGCTGCAGCTGCGCGCCCCGCCCTGGGCGCCGCCGCTGCCCCGCCAGCGTGCTGGGGCTGGCGAGCGCGAGA
AAGCAGGCCCGCGCCGCCGCCCTGCTGCTGCTGCGCCTGCTGCTGCTGGGACAACGCGCCGAGATCGGGAGCGCGAGA
CACCAACGCCCAACTACACCCCGTGCCCTGGAACACCTCTGGTCCAACAAGTCCTACAACGAGAGTCTGGAGCGCGAGA
TCTCCAACTACACCACCCAGATCTACTCCCTGATCGAGGAGTCCCAGAACCAGCAGGAGAAGAACGAGCAGGACCTGCTGGCCCTGGACAAG
TGGGCCTCCCTGTGGAACTGGTTCGACATCAGCAACTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGATCGGCCTGCGCG
CATCGTGTTCGCCGTGCTGTCCATCGTGAACCGCGTGCGCCAGGGCTACTCCCCTCTGTCCATCGTGCCTGGTGCCCGCACCGAGCTGGCCCTGG
CCGACCGCGCCGTGCTGGCCGTGGAGAGGGGCGGCGAGCAGGACAGGGACAAGAGGGGCGGCGGCGGCTGCTGGGGCCTGCTGGGG
GACGACCTGCGCTCCCTGTGCCTGTTCTCCTACCACCGCCTGCGCGACTTCATCCTGATCGCGGCGCCGAGGTGCTGGGGGCGCGGGGCTCGTGGGCCTCGCTC
CTCCCTGAAGGGCCTGGGGCGCTGGGAAGTCGTGTGGGCCTGAAGTACTGTGGAACCTGCTGCTGTACTGGGGCCAGGAGGAGCTGAAGAACTCCGCCA
TCAACCTGCTGGACACCATCGCCATCCGCCAGGGGCCTGGAGCGCGCCGTGATCGGACCCGGTGCAACCTGGCCAACCTGGCCAACCTGGAC
ATCCCCCCGCGCATCCGGCCCAGGGCCTGGAGCGCGCCATCCGCCAGGGCCTGGAGCGCGCGCCTGAAC
ATCCCCCCGCGCATCCGGCCCAGGGCCTGGAGCGCGCCTGTAA
```

Fig. 52B

2003 CON_H Env.seq.opt

```
ACCCGCGGTGATGGAGACCCAGCCGGCAACTACCCCTCCCTGTGGGCGCTGGGGCACCCTGATCCTGCTGGGCATGCTGCTCGTGCTCCGCGCCGCGG
CAACCTGTGGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAAGACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACGAGA
CCGAGAAGCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACATCATCCCCAGAGAGATGGTGCTGGAGAACGTGACCGAGAACTTC
AACATGTGGGAGAACGACACATGGTGAGCAGATGCACACCGACATCATCCCCGACATCACCGCCCACCAACGCCACCAACTCCCTGTTCAACTCCCTGTGGCGTGACCCTGACCCTGAACTGCACCGACGTGAAGAACGGCACCAACGTGACCAACAACAACTCC
CCTGTGCGTGACCCTGACCCTGAACTGCACCGACGTGAAGAACGGCACCAACGTGACCAACAACAACTCC
TCAACGTGACCACCGTGACCCGTGATCCGCGACAGCAGCAAGAAGGTGCAGCTTCTACCGCCTGGTGTCCATCGACGACAACAACTCC
TACCAGTACCGCCTGATCAACTGCAACACCTCCGTGATCACCCAGGCCTGCCCCAAGGTGTCCTTCGAGCCCATCCCCATCCACTACTGCGC
CCCCGCCGGTTCGCCACCGGCTTCGCCATCCTGAAGTGCAACAACAAGAACCTTCAACGGCACCCCTGCCCAACGTGTCCACCGTGCACCCACG
GCATCCGCCCCGTGGTGTCCACCCAGCTGCTGCTGAACGGCTCCCTGGCCGAGGAGCAGGTGGTGATCCGCTCCGCCAACCTCCGACAAC
ACCAAGAACATCATCGTGCAGCTGAACAAGCCCGTGGAGATCAACTGCACCCGGCCCAACAACAACACCCGCAAGTCCATCCACCTGGGCCC
CGGCCAGGCCTTCTACGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCTCCAAGGAGAAGTGGAACAAGACCCTGC
ACCAGGTGGTGACCAAGCTGGGCGAGTTCTTCTACTGCAACAACACCCTCCGGCCTGTTCAACTCCACCTGGGCCATGGAGAACGTGACCGAG
TCCTTCAACTGCCGCGGCGAGTTCTTCTACTGCAACAACACCCTCCGGCCTGTTCAACTCCACCTGGGGCCATGGAGAACGTGACCGAG
CATCATCACCCTGCCCTGCCGCATCAAGCAGATCGTGAACATGTGGCAGGCCGTGGGCAAGGCCATGTACGCCCCCCCATCAAGGCAACATCA
CCTGCGTGTCCAACATCACCGGCCTGATCCTGACCCGTGATCCTGACCTTCGACGAGGGCAACAACACCGTGACCTTCCGCCCCGGAGGGCGGC
AACTGGCGCTCCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTTCTTCCTGGGCCGCGCCTTCTTCCTGGGCGCAGCAGTCCAACGCTGCGCC
CGAGAAGCGCGCCGCCGCCAGCGCCGTGGGCATGGGCGCCTTCTTCCTGGGCGCAGCAGTCGCAGCAGTCCAACCTGCGCGCCTACCTGGAGGCCTGCGCC
TGCAGGGCCCGCCAGCTGCTGTCCGGCATCGTGCAGCAGCAGTCGCAGCAGTCCAACCTGCGCGCCTACCTGGAGGCCTACCTGCGCGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGG
GTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGGGCGACCAGCAGCTGCTGGGCATCTGGGGCTGCTCCGG
CAAGCTGATCTGCACCACCGCCGTGCCCTGGAACTCCTCCTGGTCCAACAAGTCCCTGGACGACAACATGACCTGGATGGAGT
GGGACAGAAGCAGAGATCAACAACAATCACACCGAGGAGATCTACCGCCTGGTGGAGGTGTCCCAGAACCAGCAGGAGAAGAACGAGCAGGACCTGCTG
GCCCTGGACAAGTGGGCCTCCCTGTGGAACTGGTTCTCCATCGTGAACTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCT
GATCGGCCTGCCGCATCATCTTCGCCGTGCTGTCCATCGTGAACCGCGTGCGCCAGGGCTACTCCCCTCTGGTCGTGCCTTCTG
ACCCCCGCGGCCCCGACCCCGAGGGCATCGAGGAGGAGGGCGACCAGGACCGCGACCGCGACAGCCTGTGGTGCGCACCGGCTTCCTG
CCCCTGGTGTGGGACGACCTGCGCCGCGGCAGGCCCGTGGACAACCTGCCGAGAGTCTGCGGAGAGCTGAAGAACCTGCGCGCCATCAACCTGC
GCTGGGGCTGGGAGATCCTGAAGTACTGGGGCAACCTGCTGCAGTACTGGGGCCAGGAGCTGAAGAACTCCCGCCACCATCAACCTGC
TGAACACCACCGCCATCGCCGTGGCCGAGGGCACCGACCGCGTGATCGAGATCGTGCAGCGGGCCTGGCCGCCGGCCATCCTGCACATCCCCCGC
CGCATCCGCCAGGGCTTCGAGCGCGCACCCTGCTGTAA
```

Fig. 53A

2003 CON_01_AE Env

```
MRVKETQMNWPNLWKWGTLILGLVIICSASDNLWVTVYGVPVWRDADTTLFCASDAKAHETEVHNVWATHACVPTDPNPQEIHLENVTENF
NMWKNNMVEQMQEDVISLWDQSLKPCVKLTPLCVTLNCTNANLTNVNNITNVSNIIGNITNEVRNCSFNMTTELRDKKQKVHALFYKLDIVQ
IEDNNSYRLINCNTSVIKQACPKISFDPIPIHYCTPAGYAILKCNDKNFNGTGPCKNVSSVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSEN
LTNNAKTIIVHLNKSVEINCTRPSNNTRTSITIGPGQVFYRTGDIIGDIRKAYCEINGTKWNEVLKQVTEKLKEHFNNKTIIFQPPSGGDLE
ITMHHENCRGEFFYCNTTKLFNNTCIGNETMEGCNGTIILPCKIKQIINMWQGAGQAMYAPPISGRINCVSNITGILLTRDGGANNTNETFR
PGGGNIKDNWRSELYKYKVVQIEPLGIAPTRAKRRVVEREKRAVGIGAMIFGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLRAIEA
QQHLLQLTVWGIKQLQARVLAVERYLKDQKFLGLWGCSGKLICTTAVPWNSTWSNRSFEEIWNNMTWIEWEREISNYTNQIYEILTESQNQQ
DRNEKDLLELDKWASLWNWFDITNWLWYIKIFIMIVGGLIGLRIIFAVLSIVNRVRQGYSPLSFQTPTHHQREPDRPERIEEGGGEQGRDRS
VRLVSGFLALAWDDLRSLCLFSYHRLRDFLLIAARTVELLGHSSLKGLRRGWEGLKYLGNLLLYWGQELKISAISLLDATAIAVAGWTDRVI
EVAQGAWRAILHIPRRIRQGLERALL$
```

Fig. 54A

2003 CON_02_AG Env

```
MRVVGIQKNYPLLWRWGMIIFWMIICNAENLWVTVYGVPVWRDAETTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEIHLENVTENF
MWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLDCHNNITNSNTTNNAGEIKNCSFNMTTELRDKKQKVYALFYRLDVVQINKNNSQYR
LINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCNDKEFNGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIVIRSENITNNAKTI
IVQLVKPVKINCTRPNNNTRKSVRIGPGQTFYATGDIIGDIRQAHCNVSRTKWNNTLQQVATQLRKYFNKTIIFANPSGGDLEITTHSFNCG
GEFFYCNTSELFNSTWNNNSTWNNTEKCITLQCRIKQIVNMWQKVGQAMYAPPIQGVIRCESNITGLLLTRDGGNNNSTNETFRPGGGDMRDNW
RSELYKYKVVKIEPLGVAPTRAKRRVVEREKRAVGLGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLRAIEAQQHLLKLTVW
GIKQLQARVLALERYLKDQQLLGIWGCSGKLICTTTVPWNSSWSNKTYNDIWDNMTWLQWDKEISNYTDIIYNLIEESQNQQEKNEQDLLAL
DKWASLWNWFDITNWLWYIKIFIMIVGGLIGLRIVFAVLTIINRVRQGYSPLSFQTLTHHQREPDRPERIEEGGGEQDRDRSVRLVSGFLAL
AWDDLRSLCLFSYHRLRDFVLIAARTVELLGHSSLKGLRLGWEALKYLGNLLSYWGQELKNSAINLLDTIAIAVANWTDRVIEIGQRAGRAI
LNIPRRIRQGLERALL$
```

Fig. 53B

2003 CON_01_AE Env.seq.opt

ATGCGCGTGAAGGAGACCCAGATGAACTGGCCCAACCTGTGGAAGTGGGGCACCCTGATCATCTGCTCCGCCCTCCGA
CAACCTGTGGGGTGACCGGTGTACTACGGCGTGCCCGTGTGGCGCGACGCCACCACCCTGTTCTGCGCCTCAAGCCCACGAGA
CCGAGGTGCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAACCCCAGGAGATCCCCTGTGGGACCAGTGAACTTC
AACATGTGGAAGAACAACATGGTGGAGCAGATGCACCAAGGACCTGATCTCCCTGTGGGACCAACATCACCAACGTGTCCAACATCACCAACG
CCTGTGCTGACCCTGAACCTGACCAACGCCAACCTGACAACGTGTCCAACATCACCAACGGTCAACATCAGGCAACATCACCAACG
AGGTGCGCAACTGCTCTTCAACATGACCAACGAGCTGCGCGACAAGAAGGTGCACGCCCTGCACCATCGTGCAG
ATCGAGGACAACAACTCCTACCGCCTGATCAACTGCAACACCTCCGTGATCAACGGCAAGAGATCTCCTTCGACCCCATCCCCAT
CCACTACTGCACCCCCCCGCGGCTACGCCATCGTGAACGGCTAGCAAGAACTTCAACGGCACCTGCCTCCTCCGAGAAC
AGTGCACCCACGGCATCAAGCCCGTGGTGTTCTACCGCACCGGCGACATCATCGGCGACATCCGCCTCCAGAGATCAACGGCACCTCCAT
CTGACCAACGCCAAGACCATCAACGTGCACCTGCAACAAGTCCGTGGAGATCAACCGCGACATCATCTTCAACAACAACACCGCCACCTCCAT
CACCATCGGCCAGGTGTTCTACCGCACCCGCGAGAAGCTGAAGGCGGGCAAGGCCGCCCCCCTACTGCCGGCCGGCCACCTGGAG
ACGAGGTGCTGAAGCAGGTGCCCCGGCATCATGTTCTACTGCAAGATCAAGCAGTTCATCAACAAGCACCTGTTCAACAACAAGGCGCCGCGCCCCC
ATCACCATGCCAACGGCACCATCATCGCCCTCTGCCCGAGTTCTTCTACTGCAACATCATCAACAGTCCGTGTTGGCAGCATCCTGCTGCGCCCCCCC
GGAGGGCTGCAACGGCACCACTTCAACTGCCGCGGCAACATCATCGCCCTGTCCAACATCACCGGCCTCCGAGCTGTACGCCCCCC
CCATCTCCGGCGGCGGCATCAACTGCGCGCTCCGAGCTGTACAAGTACACCGGCCATCGAGCGCCCCACCCG
CGCCGGCGGCAACATCAAGGACAACTGGCGGCGAGAAGCGCCGGCGCCCAGCTGCTGTCCCGGCCATCGAGCGCCGCGGCCTGCGCGCCATCGAGGCC
CAGCAGCACCTGCTGCAGGTGTGGGGCATCAAGCAGCTGCAGGCCCGGCGGCCTGCTGGAGGTGCGGCTACCTGAAGGACCAGAAGTT
CCTGGGCCTGTGGGGCTGCTCCGGCAAGATCATCTGCACCACCGCCGTGCCCTGGAACTCCAACCAGATCCTTCGAGGAGATCT
GGAACAACATGACCTGGATGGAGTGGGAGCGCGAGATCGAGAACTACACCGACATCTACGAGAGTCCAGAACCAGCAG
GACCGCAACGAAGGACCTGCTGGAGCTGGACAAGTGGGCCTCCATCATCTTCGCCGCCTGCCTGGGCTCTGTGGTACATCAAGAT
CTTCATCATGATCGTGGGCGGCCCCACCCCCGTGTCCCGGCTTCCTGGGCCTGGGACCGCCTCCTCCGGACCTTCATCCT
GTGCGCCGGCCTGGTGTGCGGCCGCCACCCCACAGGCCTGGGGCTGGTGGGAGCTGCCACCGCCTGGAACCTGC
GATCTCATGATGGTCCAGACCCCCGTGTCCCGGCCACCGTGCTCCCCTGGGCCACTCCGCCCATCTCCATCCT
TGCTGTACTGGGGCCAGGGCCTGGGCAGAGCTGAAGATCCTCCCGCGCCGCCATCCGCCTGGACGGCCTGGCCGCACCGGCGTGATC
GAGGTGGCCCAGGGCCTGGGCCGCGCATCCCGCCCATCCGCACATCCTGCACATCCCGCCAGGGCCTGGAGGCGCGGCCGCCCTGCTGTAA

Fig. 54B

2003 CON_02_AG Env.seq.opt

ATGGCGGTGATGGGCATCCAGAAGAAGAACTACCCCCTGCTGTGTGGCCTGGGGCATGATCATCTTCTGGATCATGATCATCTGCAACGCCGAGAA
CCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGCGCGAGACCGCCACCACCCTGTTCTGCGCCTCCGACGCCCAAGGCCTACGACACCG
AGGTGCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCCAACCCCCAGGAGATCCAACTGTCCCTGGAGAACGTGACCGAGAACTTCAAC
ATGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTCCCTGAAGCCCCTGAAGCTGACCCCCCT
GTGCGTGACCCTGGACTGCCGCGACAACATCACCAACAACAACAACGCCGGCGAGATCAAGAACTGCTCCTTCAACATGA
CCACCGAGCTGCGCGACAAGAAGCAGAAGGTGTACGCCCTGTTCTACGCCCTGGACGTGGTGCAGATCAACAAGAACAACTCCCAGTACCGC
CTGATCAACTGCAACACCTCCGCCATCACCCAGGCCTGCCCCAAGGTGTCCTTCGAGCCCATCCCCATCCACTACTGCGCACCCGGCATCAAGCCCG
CGCCATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACCGGCTCCCTGGCCGGAGGAGGAGTTCAACAACACGCCCCAACACCGCCAGACCATC
TGGTGTCCACCCAGCTGCTGCTGAACGGCCTGCCGTGAAGATCGTGAACAACACACCCCGCACCCCGGCACCTT
ATCGTGCAGGTGGTGAAGCCCGTGACGTGTACAAGTAGTACAACACACCCGGCCACCGTGTCCCGCCGACCAGAGTCCGTGCCCACCAAGTGGAACAACCCCTGCACGCAGGTGGCCA
CTACGCCACCGGCGACATCATCGGCGACATCTGCCACCCGACCAACATCATCTTCGCACCAAGAGACCATCTTCAACAAGATTCTTGCAACTCCTTCAACTGCGGC
CCCAGCTGCGCAAGTACTTCAACAAGACCATCATCTTCGCAACTCCTCCGAGCTGTTCAACAACACCTCCGAGAATCCAACACACCCAGGGCGTGATCCGCCTGCAGTG
GGCGAGTTCTTCTACTGCAACACCTCCGAGCTGTTCAACTGGTGGCAGAAGGTGGCAACAACAAGTGCATCACCCTGCCATCCCCATCACCGGCCATGTGAACGTGATCCGCCGCGGCG
CCGGCATCAAGCAGATCGTGAACCGCCTGCTGCTGAACCATGTGGCAGAAGGTGGGCAACAACAACTCCACCAAGAGACCTTCCGCCGGCGGCGGCGACAACTGG
TCACCGGCCTGCTGCTGACCCGCGACGGCGGCAACAACAACTCCACCAAGAGACCTTCCGCCCGGCCAAGCGCGCCAGCCGGCGAGAA
CGCTCCGAGCTGTACAAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCAAGCGCGCCTCCATCAGCCCTGACCGTGCAGG
GCGCGCGTGGGCCTGGGCCGTGTTCCTGGGCGTCCCGGCATCGTGCAGCAGTCAACCTGGCCCTGGAGGCCCATCGAGGCCCAGCTGCTGGGCCGTGTGG
CCGCCCAGCTGCTGTCCGGCATCGTGCAGCAGCAGTCAACCTGGCCCTGCTGGGCGTGCTGCTGCTGCTCCGGCAAGCT
GATCTGCACCACCGCCGTGCCCTGCCCGACATCATCATCTACAACCTGGCTGGGAACTCCTCCATCTACAACTGGCTGCAGGAGAGTCCGATCAGGCCGAGAGCCTGCTGGGGGACA
AGGAGATCTCCAACTCCAACCTGTGTGGAACTGGTTCGACATCACCAACTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGATCGG
GACAAGTGGGCCGTCCCCTGTGTTGCGCCACCCCGAGGCGGGCGGCGAGAGCGACCGCCGCGACCCCGCCTTCCTGCCCCTGATCGGCCGCCACCAGC
CCTGCGCAACGTCGTTCGCCCCGACCGGCCATCATCACACCGCCTGCTGCCCGGAGCTGTGCTGCGCCCTCCTGGCCTTCCTGCCCCTG
GCGAGCCCGACGCCGTGCCTCCCGAGGCGGCATCGAGGAGGGCGGCGGCGAGAGCGACCGCCGGCGCCCTGCTGCGCCCGCCGGAGCTGTGGG
GCCTGGGACGACGCTGCGCGCTCCCGGCGTCGTTCTCCGGCTGGGAGAGCCTGGGCGGGCTGCTGTCCTACTGGGGCCTGGAGCTGCTGGGG
CCACTCGTCCTCCCTGAAGGGGCCTGGGAGGCCGTGGCCATCGCCGCCGCTGCTGGACACACCGGCCCTCAACTGACCGGCAACCGGCGCCGCCCCGGCCATC
CTGAACATCCCCGCCGGCATCCGCCCAGGGCCCTGGAGCGGCCGCCGCCCCGCCATC·

SUBSTITUTE SHEET (RULE 26)

Fig. 55A

2003 CON 03 AB Env

MRVKEIRKHLWRWGTLFLGMLMICSATENLWVTVYYGVPVWKEATTLFCASDAKAYSKEVHNVWATYACVPTDPSPQEIPLENVTENFNMG
KNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLKKNVTSNTSSIKMMEMKNCSFNITTDLRDKVKKEYALFYKLDVVQIDNDSYRL
ISCNTSVVTQACPKISFEPIPIHYCAPAGFAILKCNDKKFENGTGPCTNVSTVQCTHGIKPVVSTQLLLNGSLAEEEVIRSVNFTDNTKTII
VQLKEPVEINCTRPNNNTRKGIHIGPGRAFYATGDIIGDIRQAHCNISITKWNNTLKQIVIKLRKQFGNKTIVFNQSSGGDPEIVMHSFNCG
GEFFYCNTTKLFNSTWNGTEELNNTEGDIVTLPCRIKQIINMWQEVGKAMYAPPIAGQIRCSSNITGLLLTRDGGNQSNVTEIFRPGGGDMR
DNWRSELYKYKVVKIEPLGVAPTKAKRRVVQREKRAVGIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQNNLLRAIEAQQHLLQL
TVWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTAVPWNTSWSNKSLDEIWNNMTWMEWEREINNYTGLIYNLIEESQNQQEKNEQEI
LALDKWASLWNWFDISKWLWYIKIFIMIVGGLVGLRIIFAVLSIVNRVRQGYSPLSFQTRLPTQRGPDRPEGIEEEGGERDRTSIRLVNGF
LALIWDDLRSLCLFIYHHLRDLLLIAARIVELLGRRGWEALKYWWNLLQYWIQELKSSAINLIDTIAIAVAEGTDRVIEIGQRFCRAIRNIP
RRIRQGAEKALQ$

Fig. 56A

2003 CON 04 CPX Env

MRVMGIQRNYPHLMEWGTLILGLVIICSASKNLWVTVYYGVPVWRDAETTPFCASDAKAYDKEVHNIWATHACVPTDPNPQEIALKNVTENF
NMWKNNMVEQMHEDIISLWDEGLKPCVKLTPLCVALNCSNATINNSTKTNSTEEIKNCSFNITTEIRDKKKKEYALFYRLDIVPINDSANNN
SINSEYMLINCNASTIKQACPKVTFEPIPIHYCAPAGFAILKCNDKNFTGLGPCTNVSSVQCTHGIKPVVSTQLLLNGSLATEGVVIRSKNF
TDNTKNIIVQLAKAVKINCTRPNNNTRKSVHIGPGQTWYATGEIIGDIRQAHCNISGNDWNETLQKIVEELRKHFPNKTIIFAPSAGGDLEI
TTHSFNCGGEFFYCNTSELFNSTYMNSTNSTTINKTITTLPCRIKQIVSMWQEVGQAMYAPPIAGSINCSSDITGILLTRDGGNNNTNNETFR
PGGGDMRDNWRSELYKYKVVKIEPVGVAPTRARRRVVQREKRAVGIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLRAIEA
QQHLLRLTVWGIKQLQARVLALESYLKDQQLLGIWGCSGKLICTTNVPWNSSWSNKSYNDIWDNMTWLQWDKEINNYTQIIYELLEESQNQQ
EKNEQDLLALDKWANLWNWFNISNWLWYIKIFIMIVGGLIGLRIIFAVLSIVNRVRQGYSPLSLQTLIPTTQRGPDRPEGTEEEGGEQDRSR
SIRLVNGFLPLIWDDLRNLCLFSYRHLRNLLLIVARTVELLGIRGWEALKYLWNLLLYWGQELRNSAINLLDTTAIAVAEGTDRIIEAVQRA
CRAIRNIPRRIRQGLERALL$

Fig. 55B

2003_CON_03_AB Env.seq.opt

ATGGCGCGTGAAGGAGAGATCCGCAAGCACCTGTGGCGCTGGGGCTGGGGCATGCTGGGCATGCTGATGATCTGCTCCGCCACCGAGAACCTGTG
GGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCACCACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACTCCAAGGAGGTGC
ACAACGTGTGGGCCACCTACGCCTGCGTGCCCACCGACCCCTCCCCCCAGGAGATCCCCCTGGAGAACGTGACCGAGAACTTCAACATGGGC
AAGAACAACATGGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTCCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGT
GACCCTGAACTGCACCGACCTGGCCGACAAGGTGAAGAAGAACGTGACCTCCAACAACACCTCCTCTACAAGCTGGACGTCGTGCAGATGAGGAACTGCTCCTTCAACA
TCACCACCGACCTGCGCGACAAGGTGAAGAAGGTGAAGATCTCTTACAAGCTGGACGTGGTGCAGATCGACAACGACTCCTACCGCCTG
ATCTCCTGCAACACCTCCGTGATCACCCAGGCCTGCCCCAAGATCTCCTTCGAGCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTTCGC
CATCCTGAAGTGCAACGACAAGAAGTTCAACGGCACCGGCCCCTGCACCAACGTGTCCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGG
TGTCCAGCTGCTGCAACGGCTCCCTGGCCGAGGAGGTGGTGATCCGCAGCGAGAACTTCACCGACAACGCCAAGACCATCATC
GTGCAGCTGAAGGAGCCCGTGGAGATCAACTGCACCCGCCCCAACAACAACACCCGCAAGGGCATCCACATCGGCCCGGGCCCTTCTA
CGCCACCGGCGACATCATCGGCGACATCCGCCAGGCCCACTGCAACATCTCCAGTGAACAACAACATCTCCGGCGACCCCGAGATCGTGATCA
AGCTGCGCGAAGCAGTTCGGCAACAAGACCATCGTGTTCAACCAGTCCTCCGGCGGCGACCCCGAGATCGTGATGCACTCCTTCAACTGCGGC
GGCGAGTTCTTCTACTGCAACACCTCCAAGCTGTTCAACTCCACCTGGAACGGCAACGGCACCTGGAACGGCGGCACCATCGTGAC
CCTGCCCTGCCGCATCAAGCAGATCATCAACATGTGGCAGGAGGTGGGCAAGGCCATGTACGCCCCCCCATCGGCGGCCAGATCCGCTGCT
CCTCCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAACAGCAGTCGCAACCGTGGGCGCGCGGCGCCAAGCGCCGGTGGTGCA
GACAACTGGCGCTCCGAGCTGTACAAGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAGGGCCAAGCGCCGCCGTGGTGCA
GCGCGAGAAGCGCGCCGTGGGGCATCGGGCCGTGTCCGGCATCGTGCCGGCATCCGTGCTGCGCGGCATCGTGGGCTGCTG
CCGTGCAGGCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGCGGCCGTGGGGCCGCGGCCTGGGCATCTGGGGCTGCTC
ACCGTGGGGGCATCAACTGCTGCAGCAGCGGCCGTGTCCGGCCCGCCCTGGAACACCTCCTGGTCCAACAAGTCCCTGGACGAGATCTGGGATGG
CGGCAAGCTGATCTGCACCACCGCGGTCGACAACAACTACACCGGCCTGATCCACTCCTACAACGACCTGGAGAACGGCAGCGAGATC
AGTGGGAGCGCCTGAACGAGTGGGCCGCCTCCCTGTGGAACTGGTTCGACATCTCCAAGTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGG
CTGGTGGCCCTGGCCCTGCTGCGCATCATCTTCGCCGTGCTGTCCATCGTGAACCGCGTGCGCCAGGGCTACTCCCCTGTCCTTCCAGACCCGCCTGC
CCACCCCAGCGGCCCGCGCGCCCCCAGGCGCCCCGCCGGCATCGAGGAGGAGGGCGGCGAGCGCGACACCACCTGCCTGCTGCGCAACGGCTTC
CTGGCCCTGATCTGGGACGACCTGCGCTCCCTGTGCCTGTTCATCTACCACCGCCTGCGCGACCTGCTGCTGATCGTGACCCGCATCGTGGA
GCTGCTGGGCCGCCGGGCCCGCCTGGGAGGCCCTGAAGTACTGGTGGAACCTGCTGCAGTACTGGGGCCAGGAGCTGCGCAACTCCGGCGTCGGATGG
TGATCGAAGTCGCCGCCATCGCCGCCGGCCTGAAGCTGGGCCGCCGTTCTGCGCGGCCGAGCTGGAGATCGAGAGGACGCCGCCGGCCCATCAACC
CGCCGCATCCGGCCATCCTGCCCGCCGTTCTGCGCCGCCGACCGCCGCCGCCAGCGCCGCCTGCCGCCGCCATCCGCCAACATCCCC
CGCCGCATCCGGCCATCCTGCCCGCCGTTCTGCGCCGCCGACCGCCGCCGCCAGCGCCGCCTGCCGCCGCCATCCGCCAACATCCCC
CTGGCCCCTGATCTGGGACGACCTGCGCTCCCTGTGCCTGTTCATCTACCACCGCCTGCGCGACCTGCTGATCGGATCGAGAGCTGCAGTACTGGTGGAACCTGCTGC
AGTCGGGACACCGGCCTGCCGCCGCGTCCGATGATCGAGAGCTGCAGCGCCGCCGTGATCGAGAGCTGCAGTACTGGGCGGC
GCTGCTGGGCCGCCGGGCCCGCCTGGGAGGCCCTGAAGTACTGGTGGAACCTGCTGCAGTACTGGGGCCAGGAGCTGCGAA
TGATCGAAGTCGCCGCCATCGCCGCCGGCCTGAAGCTGGGCCGCCGTTCTGCGCGGCCGAGCTGGAGATCGAGAGGACGCCGCCGGC
CGCCGCATCCGGCCATCCTGCCCGCCGTTCTGCGCCGCCGACCGCCGCCGCCAGCGCCGCCTGCCGCCGCCATCCGCCAACATCCCC
GCTGCTGGGCCGCCGGGCCCGCCTGGGAGGCCCTGAAGTACTGGTGGAACCTGCTGC
AGTAA

Fig. 56B

2003 CON 04 CPX Env.seq.opt

ATGCGCGTGATGGGTGACCAGCGCAACTACCCCCACCTGTGGGGAGTGGGGCACCTGTGGGCCTGGTGATCATCTGCTCCGCCTCCAA
GAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGCGCGAGACCACCCCCTTCTGCGCCTCCAAGGCCTACGACA
AGGAGGTGCACAACATCTGGGCCACCCACGCCTGCGTGCCCACCGACCCCCAACCCCCAGAGATCGCCTGAAGAACGTGACCGAGAACTTC
AACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCAACTCCACCAAGACCAACAACTCCACCGAGGAGAATCAAGAACTGCTCCTTCAACA
CCTGTGCGTGGCCCTGAACTGCTCCGACACCGAGATCCGCCACCATCAACAACTCCACCGAGAACTCCACGACCAACCTCCACCAAGACCAACAACTCCACCGAGGAGAATCAAGAACTGCTCCTTCAACA
TCAACACCGAGATCCGCGACAAGAAGAAGGAGTACGCCCTTCTACCGCCTGGACATCGTGCCCATCAACGACTCCGCCAACAACAAC
CTACTGCGCCCCGCCGGCTTCGCGGTCCCGTGGTGTCCAACGGCTCTGCTGAACGGCGTGGCCAAGTGCCCTGCCCAAGTGACCTTCGAGCCCATCCCATCCA
GCACCACCACGGCATCAAGCCCGTGGTGTCCAACGGCGTGCAGCTGCTGCTGAACGGCGTGGCCAAGTGCCCTGCCCCACCCGCGAGAACTTC
ACCGACAACAACCAAGAACACCAAGAACATCATCGTGCAGCTGAACGGCGTGGAAGATCAACTGCCGCCGTGGTGATCCGCTCCAAGAACTTCGCA
CATCGGCCCCGGCCAGACCTGGTACTGCCCACCGGACTCGGGCGAGATCATCGGCGACATCCGCCAGGCCCACTGCAACGTGTCCGTGCA
AGACCCTGCAGAAGATCGTGGAGGAGCTGCGCAAGCACCTTCGCCCCCCCCTCGCCCCGCGGCGACCTGGAGATC
ACCACCACCACTCCTTCAACTGCGCCGGCGAGTTCTTCTACTGCAACACCTCCGAGCTGTTCAACTCCACCTACTGTGGGAGGAGTGGCCACCATCAACCTCCAC
CACCATCAACAAGACCAACCTGCCCTGCCGCATCAAGCAGATCGTGTCCATGTGGGCAGGTGGCCAGGCCCATGGGCCAACAACAACAACAACGAGACCTTCGC
TCGCCGGCTCCATCATCAGCCCTGCCACCCCCTCCTCCGACAACTGGCGCTCCGAGCTGTACAAGTACAAGGTGGTGAAGATCAACTGCCGCGGCGAC
CCCGGCGGCGGCGACATGCGCGACAACTGGCGCTCCGAGCTGTGTACAAGTACAAGGTGGTGAAGATCAACTGCCGCGGCGAC
CGCCGCGCGCCGCCGCCCGCGAGAAGCGCGCCGTGGGCATCGGCCGCGTGTTCCTGGGCTTCCTGGGCGCCGCCGGCTCCACCATGG
GCGCCGCCGGCTCCATCATCACCCTGACCGTGCAGGCCCGCCAGCTGCTGTCCGGCATCGTCCAGCAGCAGAACAACCTGCTGCGCGCCATCGAGGCC
CAGCAGCACCTGCTGCTGGGCGCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGGGCAGGACCAGCAGCT
GCTGGGCATCTGGGGCTGCTCCGGCAAGCTGATCTGCACCACCACCGTGCCCTGGAACTCCACCACCTGGGACAACAAGTCCTACAACGAACATCT
GGGACAACAATGACCTGGGCAACAAGGAGATCAACAACTACACCGATCATCTACGAGCTGCTGGAGGAGTCCCAGAACCAGCAG
CTTCATCATGATCGTGGGCGGCCTGATCGGCCTGCGCATCGTGTTCGCCGTGCTCGTGGGCATCGGCATCGGCACCGATCGCCGCGATCGCCGTGCTGCTGGGCATCAGCCCGGCGGGCGGGCCGCCGGCTCCATCATCAGCCCCGGCGCAGCCCGCCGGCTCCATCATCAGCCCCGGCGGCGCAGCCCGCCGGCTCCATCATCAGCCCCGGCGGCGCAGCCCGCCGGCTCCATCATCAGCCCCGGCGGCGCAGCCCGCCGGCTCCATCATCAGCCCCGGCGGCGCAGCCCGCCGGCTCCATCATCAGCCCCGGCGGCGCAGCCCGCCGGCTCCATCATCAGCCCCGGCGGCG
TGTCCCTGCAGACCCTGGTCGCCCTGATCGCCGTGGAACGGCCTGAACTGTCCTGTGTTCTGCATCGTGAACCGGCGTGCTGCTACTCCCCCC
TCCATCCCCGCCCTGGTGAACGGCTTCCTGCCCTTCCTGTGCGAGAGCTGCTGGAGCTGTGGGGCATCAACCGCGTGGGGCAGGCGCAGACCGGCTCCCCGC
GCTGATCGTGGCCCGCCATCGTGGGAGCTGCTGGGACACCGCCATCAGCCGCGAGTACGCGATCCTGGTGAAGTACCTGTCTGCTGTACTGGGCCCAGG
AGCTGCGCAACTCCGCCGTGAGCCTCCTGAACGCCCTGGCCATCGCCGTGGGCGAGGGCACCGACCGCGTGATCGAGCTGGCCCAACCTGGCT
TGCCGCGCCATCCGCCAACATCCCCGCCCGGCATCCGGCCAGGGCCTGGAGCGCGCCCTGGCTGTGA

Fig. 57A

2003 CON_06_CPX Env

MRVKGIQKNWQHLWKWGTLILGLVIICSASNNMWVTVYYGVPAWEDADTILFCASDAKAYSAEKHNVWATHACVPTDPNPQEIALENVTENF
NMWKNHMVEQMHEDIISLWDESLKPCVKLTPLCVTLNCTNVTKNNNTKIMGREEIKNCSFNVTTEIRDKKKEYALFYRLDVVPIDDNNNSY
RLINCNASTIKQACPKVSFEPIPIHYCAPAGFAILKCRDKNFNGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIKSENLTDNTKT
IIVQLNKSVEIRCTRPNNNTRKSISFGPGQAFYATGDIIGDIRQAHCNVSRTDWNNMLQNVTAKLKELFNKNITFNSSAGGDLEITTHSFNC
GGEFFYCNTSQLFNSTRPNETNTITLPCKIKQIVRMWQRVGQAMYAPPIAGNITCTSNITGLLLTRDGNNNDSETFRPGGGDMRDNWRSELY
KYKVVKIKPLGIAPTRARRRVVGREKRAVGLGAVFLGFLGAGSTMGAASITLTVQVRQLLSGIVQQQSNLLRAIEAQQHLLQLTVWGIKQL
QARVLAVERYLKDQQLLGIWGCSGKLICPTNVPWNASWSNKTYNEIWDNMTWIEWDREINNYTQQIYSLIEESQNQQEKNEQDLLALDKWAS
LWSWFDISNWLWYIKIFIMIVGGLIGLRIVFAVLSIVNRVRQGYSPLSLQTLIPNPTGADRPGEIEEGGGEQGRTRSIRLVNGFLALAWDDL
RSLCLFSYHRLRDFVLIAARTVETLGHRGWEILKYLGNLVCYWGQELKNSAISLLDTTAIAVANWTDRVIEVVQRVFRAFLNIPRRIRQGFE
RALL$

Fig. 58A

2003 CON_08_BC Env

MRVRGTRRNYQQWMIWGVLGFWMLMICNVEGNLWVTVYYGVPVWKEAKTTLFCASDAKAYETEVHNVWATHACVPTDPNPQEIVMENVTENF
NMWNNDMVNQMHEDVISLWDQSLKPCVKLTPLCVTLECTNVSSNGNGTYNETYNESVKEIKNCSFNATTLLRDRKKTVYALFYRLDIVPLND
ENSGKNSSEYYRLINCNTSAITQACPKVTFDPIPIHYCTPAGYAILKCNDKFNGTGQCHNVSTVQCTHGIKPVVSTQLLLNGSLAEREIII
RSENLTNNVKTIIVHLNQSVEIVCTRPNNNTRKSIRIGPGQTFYATGDIIGDIRQAHCNISKDKWYETLQRVSKKLAEHFPNKTIKFASSSG
GDLEITTHSFNCRGEFFYCNTSGLFNGTYMNGTNNSSSIITIPCRIKQIINMWQEVGRAMYAPPIEGNITCKSNITGLLLVRDGGRTESNNT
EIFRPGGGDMRNNWRNELYKYKVVEIKPLGVAPTAAKRRVVEREKRAVGLGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLR
AIEAQQHMLQLTVWGIKQLQTRVLAIERYLKDQQLQTRVLAIERYLKDQQLLGIWGCSGKLICTTAVPWNSSWSNKSQQEIWDNMTWMQWDKEISNYTNTIYRLLEDS
QNQQERNEKDLLALDSWKNLWSWFDITNWLWYIKIFIMIVGGLIGLRIIFAVLSIVNRVRQGYSPLSFQILTPNPGGPGRLGRIEEEGEQD
KTRSIRLVNGFLALAWDDLRNLCLFSYHRLRDFILLTARGVELLGRNSLRGLQRGWEALKYLGSLVQYWGLELKKSTISLVDTIAIAVAEGT
DRIINIVQGICRAIHNIPRRIRQGFEAALQ$

Fig. 57B

2003 CON_06 CPX Env.seq.opt

ATGCGCGTGAAGGGCATCCAGAAGAACTGGCAGCACCTGTGGAAGTGGGGCCACCCTGGGGCCTGGTGATCATCCTGCTCCGCCTCCAA
CAACATGTGGGTGACCGTGTACTACGGCGTGCCCGCGGCCACCGACACCATCCTGTTCTGCGCCTCCGACGCGCCAAGGCCTACTCCG
CCGAGAAGCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACATCCCCAGGAGATCGCCCTGGAGAACGTGACCGAGAACTTC
AACATGTGGAAGAACCACATGGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACGAGTCCCTGAAGCCCTGCGTGAAGCTGACCCCC
CCTGTGCTGCTGTGACCCTGACCCTGAACTGCACCGACCTGACCAAGAACAACACAACGTGACCAAGAGATCATCGGCCGCCGGCAACAACTGCTCCTTCAACG
TGACCACCGAGATCCGCGACAAGAAGAAGAAGGAGTACGCCCTGTTCTACCGCCTGGACGTGGTGCCCATCGACAACAACTCCTAC
CGCCTGATCAACTGCAACGCCCTCCACCATCAAGCAGGCCTGCCCAAGGTGTCCTTCGAGCCCATCCCCATCGGCCCCCGCCGG
CTTCGCCATCCTGAAGTGCCGCGACAAGAACTTCAACGGCACGGCCCCTGGCCGAGGAGATCATCATCAACAACGTGTCCACCGGCATCAAGC
CCGTGTCCCCAGCTGCTGCCAGCTGCGTGAACGGCTCCCTGGCCGAGATCCGCTGCTGAACAAGAATCACCTCCTTCAACTGC
ATCATCGTGCAGCTGAACAAGTCCGTGAACATCGGCGACATCATCGGCCGCTGTTCAACAACACCTCCAGCTGTTCAACTGTTACGCCCCAGCTGGCCACTGTGACCCGGCC
CTTCTACGCCACCGGCGAGATCATCGGCGACATCCTGCGTTCCTGCTGGAGATCCGCGACCATCATCGGCCCGGCAACAACACCCGGC
CCGCCAAGCTGAAGGAGTTCTTCTACTGCAACAACACCTCCCAGCTGTTCAACTGTACGCCCCCCCCGGCGGCCGCAACAACGACTCCGAGACCTTCCGCGGCATCGCCCCGGCCTCCACCATGTGGGGCCGCCAGAGCGGCGCCGTGGGCCT
GCAGATCGTGCGCGACCCGCGCATGTGGCAGCCGTGGGCCAACAACGACTCCGAGACCTTCCGCGGCATCGCCCCGGCCTCCACCATGTGGGGCCGCCAGAGCGGCGCCGTGGGCCT
TGCTGCTGACCCGCGACGGCAACAACGACTCCGAGATCAAGCCCCTGGGCATCGCCCCGGCCTTCCAACCTGCTCCGAGCTGTAC
AAGTACAAGGTGGTGAAGATCGAGCCCCTGGGCGTTCCTGGCTTCCTGGGCCGTGGCAGCAGTCAAGCCCCTGGGAGCGCCAGTGCTGCTGT
GGGCCCCGTGTTCCTGGGCTTCCTGGGCGCAGCAGTCCAACCTGCTCCACCCTGGGCATCGCCGCCTGCTGGGCCATCAAGCAGCTG
CAGGCCCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCTGCTGGAGTCCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGGCCTCC
CGTGCCCTGGAACGGCCTCCTGTCCAACAAGAGACCTACAACAAGAGGACTGGGCCAACAACATGACCTGGATGGAGTGGGACCGCGAGATCAACAACT
ACACCAACAGAGATCTACTCCCTGATCGAGGAGTCCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGCCCTCC
CTGTGGTCCTGGTTCGACATCTCCAACTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGATCGGCCTGCGCATCGTGTT
CGCCGTGCTGTCCATCGTGAACCGCGTGCGCCAGGGCTACTCCCCGCTGTCCTTCCAGACCCGCCTGCCCCGCGGCGCCCGGGCCGACCGACCTG
CCGGGCGAGATCGAGGAGGGCGGCGGAGCGGCGACCGCGACCGCGACCCCTGGCCTGGGACGACCCTGGGA
CGCTCCCTGTGCCTGTTCTCCTACCACCGCCTGCGCGACTTCGTGCTGCTGGGGCGCGAGGAGCTGAAGAACTCCGCCATCGCCCTGCTGGACGCCACCGCCATCGCCG
GATCCTGGAAGTACCTGGGCAACCTGGTGCAGTACTGGGGCCTGGGGCTGGCCAGGTGGTGCAGGCCACCACCGCCATCGCCG
TGGCCAACTGGACCGACCGCGTGATCGAGGTGGTGCAGCGCGCCTGTTCCGCGCGCTGCTGCACATCCCCGCGCGCATCCCCGGCGGCCTTCGAG
CGCGCCCTGCTGTAA

Fig. 58B

2003 CON_08_BC Env seq.opt

ATGCGCGGTGCGCGGGCACCCGCCCGCCAACTACCAGCAGTGGTGATCTCGGGCGGCGTGCTGGGCTTCTGGATGATCTGATGATCTGCAACGTGGAGGG
CAACCTGTGGTGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAAGACCACCCGTTCTGCGCCTCCGACGCCCAAGGCCTACGAGA
CCGAGGTGCACAACGACAACGACATGGTGAACATGGTGAACATGTGAACAAGCGTGCCTGCGTGATGCTGAAGCTGACCCC
AACATGTGGAACAACGACATGGTGAACCAGATGCACCAGTGTCCTGTGGGACCGTGATCTCCGTGGGACACGAGATCAACGAGACCTACAACGAGTCCGTGAAGGAGATCAAGA
CCTGTGCGTGACCCTGGAGTGCACCACCCCTGCTGCGCGACCCGCAAGAGACCGTGACGCCCGTGTTCTACCGCCTGGACATCGTGCCCCTGAACGAC
ACTGCTCCTTCAACGCCACAACCTGCTGCGCGTGCGCAAGAGACCGTGTACGCCCTGATCAACTGCAACACCTCCGCCATCACCCAGGTGAACGTGCCCAGGTGACCTTCGA
GAGAACTCCGGCAAGAACTCCTCCGAGTACTACCGCCGGCTACGCCCGCGGCATCAAGCCCGTGCGCCCAGTGTGCACCGCCAGTGCTCCCAGTGCATCATCG
CCCCATCCCCCATCCCACTACTGCGCACCCCGTGCAGTGCAACAAGTGAACGGCGCTCCTGGCCGCGAGATCCCCAACAACAACAC
TGTCCACCGTGCAGTGCAACCTGACCAACACGTGAAGACCATCATCGCCGGCATCAAGCCCGTGGAGATCCGTGGAGACCGTGCCCCAACAACAACAACAC
CGCTCCGAGAACCTGACCAACACGTGCCCCGGCCAACCTTCTACGCCACCGGCGACATCGGCGACATCCCCAACAAGACCATCAAGTTCGCTCCTCCGGC
AGGACAAGTGGTACGAGAGACCCCTGCAGCCGCGTGTCCAGCGAAGCTGGCGAGCACTTCCCAACAAGACCACCTGCTGCGCGGCGAGTTCTTCTACTGCAACACCTCCGGCGAGTTCAACGGCACCTACATGAA
GGCGACCTGGAGATCACCACCCACTCCTTCAACTGCGCGGCGAGTTCTTCTACTGCCGCGGCATCAACAAGCCGCTGCGCGAGGTGGCGCGCCATGTACG
CGGCACCAACAACTCCTCTCCTCCATCATCACCTGCAACATCCCCTGGCCTGCAAGGCGCCGCCAGTGTCAACAACACC
CCCCCCCCCATCGAGGCAACATCACCTGCAAGTCCAACATCGCCAGAGTGCTGTGCTGGCGCGCCGAGTCCAACAACAACACC
GAGATCTTCCGCCCCGGCGCCAAGCCGCGCCAAGGCGCGCCCTCCAACACTGGCGGCGACATGCGCAACAACTGGCGCAACGAGTTCGAGATCAAGAGTGGTGACGATCAAGCCCGTGTTCCTGGGCGCTTCCTGGGCGCGCCG
GGCCCCACCGGCGCCAAGCCGCGCCAGGCGCCGTGCAGGCCGTGGAGCGCGAGAAGCGCCCCCAGCGTGCAGGCCCCCAGTCGTGCGACCGTGTCCGGCATCGTGCCAGCAGTCCAACCTGCTGCGC
GCTCCACCCATGGGCGCCCAGGCGCCCTCCATCACCCTGCGTGGGAGCGCGTGGAGCTGCTGCAGGCCCGTGTCCAGGCGCATCGTGCCAGCAGTCCAACCTGCTGCGC
GCCATCGAGGCCAGCAGCTGGGCCCCAGGCGCACACATGCGCGTGGGGCATCGAGGCCAGCAGCTGGGGCCATCAAGCAGTCCGGCCGTGTGGGGCATCGAGGCGCCTACCTGAA
GGACAACCAGCAGGAGATCGGGACAACAGCATGCGCAACGAGAAGGACCTGCTGGCCAACCGCGATCTCCAACAACACCACCATCTACTGGTCCTGCTCTGGAACAAGTCCC
CAGAACCAGCAGGAGGCCAACGAGAAGGACAACCTGCTGGCCCTGATCTCCTGGAAGAACCTGTGGTTCGACATCACCAACTGGCTGTG
GTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGATCGGCCTGCGCATCGTGTTCGCCGTGTCCATCGTGAACCGCGTGCGCCAGG
GCTACTCCCCCCTGTCCTTCCAGATCCTGGTGCCGCTTCCTGGTGAACGGCCTGCGGTGCGTGAAGAAGTCCCTGGGAGCTGAAGAAGTCCCACCATCTCCCTGGTGGACTCCCTGGTGGACTGGGGCGGCCTGGGCCTGAAGTACCTGGGCGCCCTGAAGTACC
AAGGACCGCTCATCCTGCTGACCGCCCGCGGCCTGGAGCTGTGGGGCGCCTGCGAGAAGCTGGTGGGACTCCCTGGTGGGAGCTGGACTACTGGGGCATCGTGCGCCAGGGCTTCGAGGCCCTGGGCCGAGGGCACC
TGGGCTCCCTGGTGGTGCAGTACTGGGGCATCCTGCTGCGCCAGGGCTTCGAGGCCCTGGGCCGAGGGCACC
GACCGGCATCATCAACACATCGTGCCGCGCCATCAACAACATCCCCGCCATCCCGCCGATCTGCCGCGCCATCCTGCACATCCCCGAGGGGCTTCGAGGGCGCCTGCGAGGGCCTGGGCCCGCCCTGCAGTA
A

Fig. 59A

2003 CON 10 CD Env

```
MRVMGIQRNCQQWWIWGILGFWMLMICNATGNLWVTVYYGVPVWKETTTLFCASDAKAYKAEAHNIWATHACVPTDPNPQEIVLENVTENF
NMWKNGMVDQMHEDIISLWDQGLKPCVKLTPLCVTLNCSDVNATNSATNTVVAGMKNCSFNITTEIRDKKQEYALFYKLDVVQIDGSNTSY
RLINCNTSAITQACPKVTFEPIPIHYCAPAGFAILKCNDKKFENGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNAKT
IIVQLNESVTINCTRPNNNTRKSIRIGPGQTFYATGDIIGNIRQAYCNISGTEWNKTLQQVAKKLGDLLNKTTIIFKPSSGGDPEITTHTFN
CGGEFFYCNTSKLFNSSWTSNNTGNTSTITLPCRIKQIINMWQGVGKAIYAPPIAGLINCSSNITGLLLTRDGGANNSETFRPGGGDMRDNW
RSELYKYKVVKIEPLGLAPTKAKRRVVEREKRAIGLGAVFLGFLGAAGSTMGAASLTLTVQARQLLSGIVQQQNLLRAIEAQQHLLQLTVW
GIKQLQARVLAVESYLKDQQLLGIWGCSGKHICTTNVPWNSSWSNKSLEEIWDNMTWMEWEREIDNYTGLIYSLIEESQNQQEKNEQELLQL
DKWASLWNWFSITNWLWYIKIFIMIVGGLIGLRIVFAVLSIVNRVRQGYSPLSFQTLLPAPRGPDRPEGIEEGGEQGRGRSIRLVNGFSAL
IWDDLRNLCLFSYHRLRDLILIATRIVELLGRRGWEAIKYLWNLLQYWIQELKNSAISLLDTTAIAVAEGTDRAIEIVQRAVRAVLNIPTRI
RQGLERALLS
```

Fig. 60A

2003 CON 11 CPX Env

```
MRVKETQRNWHNLWRWGLMIFGMLMICNATENLWVTVYYGVPVWKDADTLFCASDAKAYSTEKHNVWATHACVPTDPNPQEIPLENVTENF
NMWKNNMVEQMHEDIISLWDESLKPCVKLTPLCVTLNCTDVKNATNTVEAAEIKNCSFNITTEIRDKHKEYALFYKLDVVPINDNNNSIY
RLINCNVSTVKQACPKVTFEPIPIHYCAPAGFAILKCNDKKFENGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAEGEVRIRSENFTNNAKT
IIVQLNSSVRINCTRPNNNTRKSIHIGPGQAFYATGDIIGDIRQAHCNISRAEWNNTLQQVAKQLRENFNKTIIFNNPSGGDLEITTHSFNC
GGEFFYCNTSRLFNSTWNNDTRNDTKQMHITLPCRIKQIVNMWQRVGQAMYAPPIQGKIRCNSNITGLLLTRDGGNNTNETFRPTGGDMRD
NWRSELYKYKVVEIKPLGVAPTRAKRRVVEREKRAVGIGAVLLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLKAIEAQQHLLKLT
VWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTNVPWNFSWSNKSYDEIWDNMTWIEWEREINNYTQTIYTLLEESQNQQEKNEQDLL
ALDKWASLMNWFDISNWLWYIKIFIMIVGGLIGLRIIFAVLSIVNRCRQGYSPLSFQTLPNHKEADRPGGIEEGGGEQDRTRSIRLVSGFL
ALAWDDLRNLCLFSYHRLRDFILIAARIVETLGRRGWEILKYLGNLAQYWGQELKNSAISLLNATAIAVAEGTDRIIEVVHRVLRAILHIPR
RIRQGFERALLS
```

Fig. 59B

2003 CON 10 CD Env.seq.opt

ATGGCCGTGATGGCCATCCAGCGCAACTGCCAGCAGTGGTGGATCTGGGGCATCCTGGGCTTCTGATGCTGATGATCTGCAACGCCACCGG
CAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGACCACCACCACCCTGTTCTGCGCCTCCGACGCAACGTTGACGGACTTC
CCGAGGCCCACAACATCTGGGCACACGCCTGCGTGCCGCCCACGAGGACATCATCTCCGACGACGAGCTGAAGGCCCTGGGACGTGCGCTGAAGCTGACCCC
AACATGTGGAAGAACGGCATGGTGGACCAGATGCACGAGGACATCAACTGCGAAAGCCCTGGGTGGCCGCATGAGAACTGCTCCTTCAACA
CCTGTGTGCGTGACCCTGAACTGCTCCGACGTGAACGCCACCAACTGCGAAGATGAAGAACTGCTCCAACACCTCCTAC
TCACCACCGAGATCCGCGACAAGAGAAGCAGGAGTACGCCCTGCCCCAAGGTGGTGTCAGATCGACGGCTCCAACACCTCCATCCCACTACTGCGCCCCGCGG
CGCCTGATCAACTGCAACACCTCCGCACGAGTTCAACGGCACCCAGGCCTGCCCATCCCATCCCCGTGCCACCTTCCACCACCCCAAGACC
CTTCGCCATCCTGAAGTGCAACGACACGCTGCTGAACGGCTCCCTGGCCGAGGAGATCATCATCAACAACACCCGAGAAGTCCATCCGCGCCCAGAGACC
CCGTGGTGGTCCAGCTGAACGAGTCCGTGACCATCGGCCAACATCGGCCCGTGGAACCTGGAACAAGAACTCCGGCCACCGTGGACCGAGGAGTGG
CTTCTACGCCACCGGCGACCTGCTGAACGTGCTGCCACCATCTTCAAGCCCTCCTCCTGGACCTGTTTCAACGCCTGAGTGGGGCCAAGCCCATCCTGC
TGCGGCGGCGAGTTCTTCTACTGCAACATCAACATGTGGCAGGCCGTGGGGCGCGCCAACAACCTGCTGGGCCCGCGCCAAGGCCCGGGAGGCGGAGAA
GCCCTGCCGCGCATCAAGCAGATCATCAACATGTGGCAGGCCGTGGGGCGGGGCCATCGAGGGCCCATCGAGGGCCCAGCACCTGCTGGGGCCTGGTGG
CCCGCAGCTGCTGTCCGGCGGCGATCGTGCAGCAGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACCTGCTGGGGCTGGTGGTCCGGCCAAGCA
GGCCATCAAGCAGCTGCCCGCGGCCCGGGCCTGCGGAGTCCTTACCTGGAGTCCTCCAACAAGTCCCTGGAGGAGATCTGGGACAACTGGGAGTGGGGAGC
CATCTGCACCACCAAGGTGCCCTGGAACTCTCCTGGTCCAACAAGTCCCTGGAGGAGTCCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGCCTG
GCGAGATCGACAACTACACCGGCCTGATCTACTCCCTGATCGAGGAGTCCCAGAACCAGCAGGAGAAGAACGAGCAGGAGCTGCTGGCCCTG
GACAAGTGGGCCTCCCTGTGGAACTGGTTCTCCATCACCAACTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTGATCGG
CCTGCGGATCGTGTTCGCCGTGCTGTCCGTGGTGAACCGCGTGCGGCAGGGCTACTCCCCGCTGTCCTTCCAGACCCTGCTTCCGCCCCTG
GCGGCCCCCGACCGGCCATCGAGGAGGGCGGCGAGCGGGGCGCGGGGCGAGGGCGAGCGGGACCGGGAGCGGGACCGGTCCCGGCAAGCA
ATCTGGGACGACCTGCGCAACCTGCGCGCGCTGCTGGAACCTGGACGCGCCATCCTGAGCTGGAGAGCTCCTGCTGCTGGACA
CCGCCCGCCGACCTGGGGAGGCCGTGGCCCGTGGGGCGCGGCGAGATCGTGCAGCGCCTGCAGCGCGCCGTGCGCCGTGCTGCTGGG
CGCCCAGGGCCTGGGAGCGCGCGCCCTGCTGTAA

Fig. 60B

2003 CON 11 CPX Env.seq.opt

```
ATGCGCGTGAAGGAGAGACCCAGCGCAACTGGCACAACCTGTGGGCGCTGGGGCCTGATGATCTTCGGCATGCTGATGATCTGCAACGCCACCGA
GAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGACGCCGACACCACCCTGTTCTGCGCCTCCGACGCCAAGGCCTACTCCA
CCGAGAAGCACAACGTGTGGGCCACCCACGCCTGCGTCCCGACCCCAACCCCCAGGAGATCATCCTCCCTGTGGGACGAGTCCCTGAAGCCTGACCCC
AACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCATCAGCCTGTGGGATCAAGACTGCAAGTGACGTGCACAACTGAACGAACTTC
CCTGTGCGTGACCCTGAACTGCACAGATGCAAGAAACGCCACCAACACCACCGTGAGAGGGCCGCCGAGATCAAGAACTGCTCCTTCAACATCA
CCACCGAGATCAAGGACAAGAAGAAGAGTACGCCCTTGTTCTACAAGCTGGACGTGGTGCCCATCAACGACAACAACTCCATCTAC
CTTCGCGATCAATCCTGAAGTGCAACGTGTCAACGACAGAAGTTCAACGCGCTCCTGCTGCCCCAAGGTGACCTTGCAGCCCATCCCCACCGTGTCCACGTGCCAGAGAACGTGTCCGCTCCGAGAACTTCACCAACAACGCCAAGACC
CCGTGGTGTCCACCCAGCTGCTGCTGAACTGCAATCGGCGCATCAACTGCACACCCGGCCCACTGCAACATCCGCGACACTGCGGCCAGGC
ATCATCGTACGCCACCGGCGACGAGAACTTCAACAAGACCATCATCTTCAACACCCTCCGGCGGCGACCTGGAGATCACACACCCACTCCTTCAACTGC
CTTCTACGCCACCGGCGACGAGAACTCCGGCGGCGACCTGGAGATCACACACCCACTCCTTCAACTGC
CCAAGCAGCTGCGCGAGAACTTCAACAAGACCATCATCTTCAACACCCCTCCGGCGGCGACCTGGAGATCACACACCCACTCCTTCAACTGC
GGCGGCGAGTTCTTCTACTGCAACAGGACCTCCCGCGCTTCAAGCAGATCGTGAACATGTGGCAGCGCGGCAACAACACCAACAAGACCCCTGGGCGCTT
CCTGCCCCTGCCGCATCAAGCAGATCGTGAACATGTGGCAGCGCGGCAACAACACCAACAAGACCCCTGGGCGCTT
ACTCCAACATCACCGGCCTGCTGCTGACCCGGCGCGGCAACAACAACAACAACCAAGCGGCAACATGTGGCAGATCCGCGCGCA
AACTGGCGCTCCGAGCTGTACAAGTACAAGGTGTGGTGAGATCAAGCCCCTGGGCGTTCCTGGCGCCCGCCATGGGCGCGCGAC
CGAGAAGCGCCCCGGCATGGGCGCGCGGCGCCGTGGGGGCTTCCTGGGCTTCCTGCAGCAGCAGTCCAACCTGGCCGTGGAGGCG
TGCAGGCCCGCCGCCAGTGCTGCTTGGCGGCCATCGTGCCGCGCGCTGCTGCCCGTGGCCGTGGGGAGCTGACCG
GTGTGGGGGCATCAAGCAGCTGCACCACCAACGTGCCCCTGCCGGCAGGCCCATCAACCTGGCTGGGGCGCTGCTGCTGACCCTGAAGCTGACC
CAAGCTGATCTGCACCACCCCAGGTGCCCTGGAAATTCTCCTGGTTCAACAAGTCCTACAACATGACCTGGGACACCAACATGACCTGGGACACCAACATGACCTGGATCGAGT
GGGAGCGCGAGATCAACAACTACACAACCTGCTGGGACGAGATCGGGGTCCCAGAACCTGGTGTACAACTGGCTGTGGTCCCCAGAAGAACGAGAGGACCTGCTG
GATCGGCCTGCTGCTGCATCATCTTCGCCGTGCTGTCCCTGGTTCAACTGGCTGTGCCTGTGTCGACATCTCTGAACCTGTGACATCTGAACCGCTGCCATCTCCCCTGGGGGCCGGCCT
ACCACAAGGAGGCGCGGCTGGGGACGAGATCCTGCGGCCGGGCGGCCCCAGACCCCTGACCCCCA
GCCCCTGCCTGGGGACGAGATCCTGCGGCCGGGCGGCCCCAGACCCCTGACCCCCA
CCTGGGCGCTGCTGCGCCAACTGTGCCGTGTTCTCCACCACCGCCGACTTCATCTGCGCCCGTGTCCGCCATCGTGGAGAC
TGAACGCCACCACCGCCGGCCCGAGGGCCACCGACCCATCATCAGGGCCAGGAGCTGAAGAACTCCGCCATCTCCCTGC
CGCATCCCGCCAGGGCTTCCGAGCGCGCCAGGGCCGCCGCCCTGCTGTAA
```

Fig. 61A

2003 CON_12_BF_Env

MRVRGMQRNWQHLGKWGLLFLGILILICNATENLWVTVYYGVPVWKEATTLFCASDAKSYEREVHNVWATHACVPTDPNPQEVDLENVTENF
DMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLNCTDANATANATKEHPEGRAGAIQNCSFNMTEVRDKQMKVQALFYRLDIVPISDN
NSNEYRLINCNTSITTQACPKVSWDPIPIHYCAPAGYAILKCNDKKFENGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEIIRSQNIS
DNAKTIIVHLNESVQINCTRPNNNTRKSIHIGPGRAFYATGDIIGDIRKAHCNVSGTQWNKTLEQVKKKLRSYFNTTIKFNSSSGGDPEITM
HSFNCRGEFFYCNTSKLFNDTVSNDTIILPCRIKQIVNMWQEVGRAMYAAPIAGNITCTSNITGLLLTRDGGHNETNKTETFRPGGGNMKDN
WRSELYKYKVVEIEPLGVAPTRAKRQVVKREKRAVGIGALFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQSNLLRAIEAQQHLLQLTV
WGIKQLQARVLAVERYLKDQQLLGLWGCSGKLICTTNVPWNSSWSNKSQEEIWENMTWMEWEKEINNYSNEIYRLIEESQNQQEKNEQELLA
LDKWASLWNWFDISNWLWYIRIFIMIVGGLIGLRIVFAVLSIVNRVRKGYSPLSLQTHIPSPREPDRPEGIEEGGGEQGKDRSVRLVNGFLA
LIWDDLRSLCLFSYHRLRDLLLIVTRIVELLGRRGWEVLKYWWNLLQYWSQELKNSAISLLNTTAIVVAEGTDRVIEALQRVGRAILNIPRR
IRQGLERALLS

Fig. 62A

2003 CON_14_BG_Env

MKAKGTQRNWQSLWKWGTLILGLIVIICSASNDLWVTVYYGVPVWKEATTLFCASDAKAYDAEVHNVWATHACVPTDPNPQEVALENVTENF
NMWENNMVDQMQEDIISLWDQSLKPCVELTPLCVTLNCTDFNNTNNTNTRNDGEGEIKNCSFNITTSLRDKIKKEYALFYNLDVVQMDND
NSSYRLTSCNTSIITQACPKVSFTPIPIHYCAPAGFVILKCNNKTFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAEEIVIRSKNFTD
NAKTIIVQLKDPIEINCTRPNNNTRKRITMGPGRVLYTTGQIIGDIRKAHCNISKTKWNNTLGQIVKKLREQFMNKTIVFQRSSGGDPEIVM
HSFNCGGEFFYCNTLQLFNSTWRSNSTWNDTTETNNTDLITLPCRIKQIVNMWQKVGKAMYAPPISGQIRCSSNITGLLLIRDGGSNNTETF
RPGGGNMKDNWRSELYKYKVVKIEPLGVAPTRAKRRVVQREKRAVGIGALLFGFLGAAGSTMGAASMTLTVQARQLLSGIVQQQNNLLRAIE
AQQHMLQLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICTTTVPWNASWSNKSLDDIWNNMTWMEWEREIDNYTGLIYTLIEQSQNQ
QERNEQELLELDKWASLWNWFNITNWLWYIKIFIMIIGGLIGLRIVFAVLSIINRVRKGYSPLSFQTLTHHQREPDRPGRIEEEGGEQDKDR
SIRLVSGFLALAWDDLRSLCLFSYHRLRDFILIAARTVELLGRSSLKGLRLGWEGLKYLWNLLLYWGRELKNSAINLLDTVAIAVANWTDRA
IEVVQRVGRAVLNIPVRIRQGLERALLS

Fig. 61B

2003 CON 12 BF Env.seq.opt

ATGGCGCGTGCGCGGCCATGCCAGCGCAACTGGCAGCACCTGGGCAAGTGGGGCCTGCTGTCTGTTCCTGGGCATCATCTGCAACGCCACCGA
GAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCACCACCACCCTGTTCTGCGCCTCCGACGCCAAGTCCTACGAGC
GCGAGGTGCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACAATCCCCAGGAGGTGCCCACCGACATCTCCCTGTGGGACCAGTCCTT
GACATGTGGAAGAACAACATGGTGGAGCAGATGCACGAGGACATCATCTCCCTGTGGGACCAGTCCCTGTGGGACCAGTCCCTGTGGGACC
CCTGTGTGCGTGACCCTGAACTGCACCGACCTGGCCCAACGCCAACGCCACCGACGACGCCAAGAGCCAGGGCCCGCGCCATCTCCAGAACT
GCTCCTTCAACATGACCACGAGTGGCGCGACGAGGTGCGCGAACAAGCAGATGAAGGTGCGCGCCCTGTTCTACCGCGCCTGGACATCTCCGACAAC
AACTCCAACGATACCGCCTGATCAACTGCAACACTCCCACCAGCAAGACCACACACCTCCAACCACAAGAAGTTCAACGACAAGAGTGCAACGAC
CTGCGCCCCCCGCGCTACGCCGTGGTGCTCAAGCTGCTGAACGACCCCGGTGTCCTGCTGAACGCTGCCGGCCGAGGAGAGTTCAACGACAAGAGTGCAGTGCA
CCCACGCCATCAAGCCCGTGGTGTCCAACGCTGCTGAACGACCCGTGTCCTGCTGAACGCTGCCGGCTCCGTGCCCAGAACATCTCC
GACAAGCGCCCGGCCGCCTTCTACGCCGCCATCATCGTGCACCTGCTGAACGACGAGTCCGTGCAGATCAACGCGCGACATCATCGGCGAACGTGTCCGGCGCACCCAGTGAACAAGA
CGGCCCGGCCGCGCCCTTCTACGCCGCGATCATCGGCGACATCATCTACTGCAACCGGCTCCTCTTCAACACCACCCTCCAAGCTGTTCAACGACAACCGTGGAACAAGA
CCCTGGAGCAGGTGAAGAAGAAGCTGCGCTCCGAGTTCTTCTACTGCAACACCACCCTCCAAGCTGTTCAACGACACCGTGTACGCCCCCGAGATCATCCTGCCCTG
CACTCCTTCAACTGCGCCGGGCGAGTTCTTCTACTGCAACACCACCCTCCAAGCTGTTCAACGCCCCATCGCGCCGGCCAACATCCTGCCCTG
CCGGCCTCCGAGCTGTACAAGTACAAGGTGGTGACGACGGCGGCCACAACGAGACCAACAAGAAGACCAACAAGAACCGGCCCCCGGCGCCAACATGAAGGACAAC
TGGCGCTCCAGATCGTGCTGCTGACCCGCGATCGGTGGGATCGAGGCCCGCCCGCCCACCACCCGCCAACATGAAGGACAAC
GAAGCGCGCCCGGCCGAGGTGGGCATCGGGCCGCCCTGTTCCTGGGCTCTTCCTGGGCTCCAGCAGCAGCCGGCCGTGGCCAACCTGCTGCAGCTGACCGGTG
AGGCCCCGCGCCAGCTGCTCCGGCATCGGCCCGCCCTGTTCCTGCGCCAGCAGCCGGCCGTGGCCAACCTGCTGCAGCTGACCGGTG
TGGGGCATCAAGCAGCTGCAGGCGCGCGTGCTGGCCGTGGAGCGCTACCTGAAGGACCAGCAGCTGCTGGGCCTGTGGGGCTGCTCCGGCAA
GCTGATCTGCACCACCACCGCCGTGCCTGGAACTGCCCTGGTCCAACAAGTCCCTGATCGAGATCTACCGCGACATCTCCAACTGGCTGGTACATCTCCAACTGGCTGGT
AGAAGGAGAGTCAACAACTACTACCTGGAAGCTGTCCCGCCTGATCGAGATCTACCGCGACATCTCCAACTGGCTGGT
CTGGACAAGTGGGCCCTCCCTGTGTGCGCCGTGTTCGCCGTGCTCTGTCTCATCTGCAACGGCGCGCGGGGAGCCTGTGGGCGGCCTGAT
CGGCCTGCGCATCGTGTTCGCCGTGCTCTGTCTCATCGTGAACCGGCGGGGCCGCAAGGCTACTCCCTCAAGGGCGCGGAGGCTACTTCCTGGCC
CCCGGCGAGCCCGACCGCGCCCCGAGGGGCATCGAGGAGGGCGAGCAAGGACCGCGGCCCCGCCGCCGCGCCATCGTGGAAACGGCTTCCTGGCC
CTGATCTGGGACGACCTGCGCGCCTCCGTGTTCCTGCTGTTCTCCTACCACCGCCTGCGCGACTTCCTCTTCCTGGAAGACCTGATCGTGGAGCTGCT
GGGCCGCCGCGGCTGGGGAGGTGCTGAAGTACTGGGCGAGGCCACCGAGGGCCACCGCGTGATCGAGGGCCCTGGGCGGTACTGGGTGAACCGCGTGATCGAGGGCCACCGCGTGATCGAGGGCCCTGGGCGG
ACACCACCACGTGTGGGGGCCGAGGGCCACCGCGTGATCGAGGGCCACCGCGTGATCGAGGGCCCTGGGCGCGTGATCCTGAACATCCCCCGCGGC
ATCCCGCCAGGGCCTGGAGCGCCTGGAGCGCGCCCTGCTGTAA

Fig. 62B

2003 CON_14_BG Env.seq.opt

```
ATGAAGGCCAAGGGCACCAGCGCAACTGGCAGTCCCTGTGGAAGTGGGCACCCTGATCCTGGTGATCATCTGCTCCGCCTCCAA
CGACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGTGGCGCCACCACCACCCCACCACCCCCAGGAGGCCACCACCACCCG
CCGAGGTGCACAACGTGTGGGCCACCCACGCGCCTGCCGGTGCGTGCCACCGACATCATCTCCCGAGCGCTGGTCTGCGCCCAAGGCCTACGACG
AACATGTGGGAGAACAACATGGTGGACCAGATGCAGGAGGACATCATCTCAACACCACCAACAACACCACCAACACCCGAGAACGTGACCCCC
CCTGTGCGTGACCCTGAACTGCACCGACTTCAACAACAACACCACCAACAGATCAAGAGGAGTACGCCCTGTTCTACAACGGCGAGATGAAGAACT
GCTCCTTCAACATCACCACCAGCATCCGCGACAAGATCCATCATCCCCGACACACCTGCTTCACCAAGGTGTCCTTCACCCCATCCCATCCCCATCCTG
AACTCCTCCTACCGCGGCTTCGTGATCCTGAAGTGCAACAACAAGACCTTCAACGGCTCCTGGCCGAGGAGGATCGTGATCCGCTCCAAGAACTTCACCGAC
CGCCCCGGCTACGTGTCTGCTGCTACACCAGCTGCTGCTGCTGGCACCCATCGACCGGCCAGATCAACTGCGACCATCCGCGGCTCCAAGGCCAGTGGGG
ACGGCATCCGCGCCGGTGCGTGTACACCAGCTTGCCGGAGCAGCAGTTCATGAACAAGACCATCGTGTTCAACAACAACACCAGCTGGGCCAAGTGGAACACCCC
TGGGCCAGATCGTGAAGAAGCTGCGCGGGCGGCGGAGTTCTTCTACTGCAACACACCACCCCAGCTGTGGAACGAATCTGGGAACGACACAC
CACCGAACGCAACACCGACCCGCGAGGTCCCTGCTCCCTGCCGATCAGTGTTCAACTCCAACAGATCGTGATCGGCCTCAAGCAGGCAGTGCCCATGTACG
CCCCCCATCTCCGGCCAGGCCAAGCTGCGTGGGGCGCCAGCATCAACTGCGTGCTCCAACAATCTACCGGCCAAGCCGTCCCGGCCTCCACCA
CGCCCGGCCGGCGAAGCGCCGCTCCATGACCCTGGTGCAACTGGCCGAGAAGCGCGCCCGCCAGCGTGCGGGCGCGTGTGTGCGCGGCTCGCGAG
TGGGCGGCCGCCTCCATGACCTGGGGGCTGCTCCGCGCGCCGAGTGGGGCTGGTGGCCCAGCATCAACTGCTGCGTGGAGAGCGCGTGACCTGGAGGACCAGCA
GCCCAGCAGCACATGCTGACCTGCTCCGGCGCGCGGTGTGGGGCATGCTGCTGGAGCGCCGAGTGGCCTCGTCCAACAAGTCCCTGACGACAA
GCTGCTGGGCATGCTGGGGCTGCTCCGGCACCTGGGAGTGGGATGGGAGCTGCTGGACAAGTGGGAGCTGGCTGCGCAAGGGCTACTCCC
CAGGAGAACGACGCGCAACTGGGCCCTGACCGTGGGGGAGCTGGACAAGTGCGCGGATCGACAACTGTTCAACATCACCAACTGGCTGTGTGGGTACATCAA
GATCTTCATCATGATCGTGGGCGGCCTGATCGGCCTGCGGATCGTGTTCGCCGTGCTGTCCATCATCAACCGCGTGCGCCAAGGGCTACTCCC
CCCTGTCCCTTCCAGACCCTGCTGCCCGCTCCTCCTGGCCTCGCTGGAGCTGCTGGGCCTCCTGGGAGGGCCTGAAGGCGCCTGAACCTGTGGAACC
CCTGCTGAAGTACTGGGGCAACTGGGGCGCCCTGAAGTACTGGTGGAACCTGCTGGCCTACTGGAGCCAACTGGGCCTGGAGGGCGGCTGGAGCGGCGCC
ATCGAGGTGGTGCAGCGCCTGGGCCGCGCATCCCGGCCATCCCCGTGCTGCTGAACATCCCGGCCTGGAGCGGCGCCCTGGAGCGGCGCCGCC
CTGGCTGCTGTACTGGGGCCGCGCAGCTGTACTGGGGCCCGGCCATCCCCGTGCTGCTGAACATCCCGGCCTGGAGCGGCGCCTGGAGCGGCGCCGCC
GCTGCTGTACTGGGGCCCGCCAGGGCCCTGGAGCGGCGCCTGGAGCGGCGCCGCC
ATCGAGGTGGTGCAGCGCCGTGGGCCCTGGAGCGGCGCCTGGAGCGGCGCCGCC
```

Centralized HIV-1 gag/nef/pol Protein and the Codon-optimized Gene Sequences

Fig. 63A

1. 2003_CON_s gag.PEP

MGARASVLSGGKLDAWEKIRLRPGGKKKYRLKHLVWASRELERFALNPGLLETSEGCQQIIEQLQPALQTGSEELRSLYNTVATLYCVHQRI
EVKDTKEALDKIEEEQNKSKQKTQQAAADTGNSSKVSQNYPIVQNLQGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDL
NTMLNTVGGHQAAMQMLKDTINEEAAEWDRLHPVHAGPIPPGQMREPRGSDIAGTTSTLQEQIGWMTSNPPIPVGEIYKRWIILGLNKIVRM
YSPVSILDIRQGPKEPFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQNANPDCKTILKALGPGATLEEMMTACQGVGGPSHKARVLAEAMS
QVTNTTIMMQRGNFKGQKRIIKCFNCGKEGHIARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSNKGRPGNFLQSRPEPTAPPAE
SFGFGEEITPSPKQEPKDKELYPLASLKSLFGNDPLSQ$

Fig. 63B

2003_CON_s gag.OPT

ATGGGCGCCCGCGCCTCCGTGCTGTCCGGCGGCAAGCTGGACGCCTGGGAGAAGATCCGCCTGCGCCTGCGCGGCGGCAAGAAGAAGTACCGCCT
GAAGCACCTGGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCCTGAACCCCGGCCTGCTGGAGACCTCCGAGGGCTGCCAGCAGATCATCG
AGCAGCTGCAGCCCGCGCTGCAGACCGGCTCCGAGGAGCTGCGCTCCCTGTACAACACCGTGGCCACCCTGTACTGCGTGCACCAGCGCATC
GAGGTGAAGGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGAACAAGTCCAAGCAGAAGACCCAGCAGGCCGCCGCCGACACCGG
CAACTCCTCCAAGGTGTCCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAGGCCATCTCCCCCCGCACCCTGAACG
CCTGGGTGAAGGTGGTGGAGGAGAAGGCCTTCTCCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCCCAGGACCTG
AACACCATGCTGAACACCGTGGGCGGCCATCCCCCCGGCCAGATGCGCGAGCCCCGCGGCTCCGACATCGCCGGCACCACCTCCACCCTGCAGGAGC
AGATCGGCTGGATGACCTCCAACCCCCCCATCCCCGTGGGCGAGATCTACAAGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATG
TACTCCCCCGTGTCCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCCTGCGCGCCGA
GCAGGCCACCCAGGACGTGAAGAACTGGATGACCGACACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCCCTGG
GCCCCGGCGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTGCTGGCCGAGGCCATGTCC
CAGGTGACCAACACCACCATCATGATGCAGCGCGGCAACTTCAAGGGCCAGAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCA
CATCGCCCGCAACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGG
CCAACTTCCTGGGCAAGATCTGGCCCTCCAACAAGGGCCGCCCCGGCAACTTCCTGCAGTCCCGCCCCGAGCCCACCGCCCCCCCGCCGAG
TCCTTCGGCTTCGGCGAGGAGATCACCCCCTCCCCCAAGCAGGAGCCCAAGGACAAGGAGCTGTACCCCCTGGCCTCCCTGAAGTCCCTGTT
CGGCAACGACCCCCTGTCCCAGTAA

Fig. 64A

2. 2003_M.GROUP.anc gag.PEP

MGARASVLSGGKLDAWEKIRLRPGGKKKYRLKHLVWASRELERFALNPGLLETAEGCQQIMGQLQPALQTGTEELRSLYNTVATLYCVHQRI
EVKDTKEALDKIEEEQNKSQKTQQAAADKGDSSQVSQNYPIVQNLQGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDL
NTMLNTVGGHQAAMQMLKDTINEEAAEWDRLHPVHAGPIPPGQMREPRGSDIAGTTSTLQEQIGWMTSNPPIPVGEIYKRWIILGLNKIVRM
YSPVSILDIRQGPKEPFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQNANPDCKTILKALGPGATLEEMTACQGVGGPGHKARVLAEAMS
QVTNANIMMQRGNFKGPRRIVKCFNCGKEGHIARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSNKGRPGNFLQSRPEPTAPPAE
SFGFGEEITPSPKQEPKDKELYPLASLKSLFGSDPLSQ$

Fig. 64B

2003_M.GROUP.anc gag.OPT

ATGGGCGCCCGCGGCCTCCGTGCTGTCCGGCGGCAAGCTGGACGCCTGGGAGAAGATCCGCCTGCGCCTCCCGGCGGCAAGAAGAAGTACCGCCT
GAAGCACCTGGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCCTGAACCCCGGCCTGCTGGAGACCGCCGAGGGCTGCCAGCAGATCATGG
GCCAGCTGCAGCCCGCGCTGCAGACCGGCACCGAGGAGCTGCGCTCCCTGTACAACACCGTGGCCACCCTGTACTGCGTGCACCAGCGCATC
GAGGTGAAGGACAAGAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGAAGACCCAGCAGGCCGCCGCCGACAAGGG
CGACTCCTCCCAGGTGTCCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAGGCCATCTCCCCCGCCACCCTGAACG
CCTGGGTGAAGGTGGTGGAGGAGAAGGCCTTCTCCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCCCAGGACCTG
AACACCATGCTGAACACCGTGGGCGGCCATCCCGCCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCT
GCACCCCGTGCACGCCGGCCCCATCCCCCCAGGCCAGATGCGCGAGCCCCGCGGCTCCGACATCGCCGGCACCACCTCCACCCTGCAGGAGC
AGATCGGCTGGATGACCTCCAACCCCCCCATCCCCGTGGGCGAGATCTACAAGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATG
TACTCCCCCGTGTCCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCCTGCGCGCCGA
GCAGGCCACCCAGGACGTGAAGAACTGGATGACCGACACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCCCTGG
GCCCCGGCGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCATCTGAAGGTGCTTCAACTGCGGCAAGGAGGGCCA
CATCGCCCGCAACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGG
CCAACTTCCTGGGCAAGATCTGGCCCTCCAACAAGGGCCGCCCCGGCAACTTCCTGCAGTCCCGCCCCGAGCCCACCGCCCCCCCGCCGAG
TCCTTCGGCTTCGGCGAGGAGATCACCCCCTCCCCCAAGCAGGAGCCCAAGGACAAGGAGCTGTACCCCCTGGCCTCCCTGAAGTCCCTGTT
CGGCTCCGACCCCCTGTCCCAGTAA

Fig. 65A

3. 2003 CON A1 gag.PEP

MGARASVLSGGKLDAWEKIRLRPGGKKKYRLKHLVWASRELERFALNPSLLETTEGCQQIMEQLQPALKTGTEELRSLYNTVATLYCVHQRI
DVKDTKEALDKIEEIQNKSKQKTQQAAADTGNSSKVSQNYPIVQNAQGQMVHQSLSPRTLNAWVKVIEEKAFSPEVIPMFSALSEGATPQDL
NMMLNIVGGHQAAMQMLKDTINEEAAEWDRLHPVHAGPIPPGQMREPRGSDIAGTTSTPQEQIGWMTGNPPIPVGDIYKRWIILGLNKIVRM
YSPVSILDIKQGPKEPFRDYVDRFFKTLRAEQATQEVKNWMTETLLVQNANPDCKSILRALGPGATLEEMMTACQGVGGPHKARVLAEAMS
QVQHTNIMMQRGNFRGQKRIKCFNCGKEGHLARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSSKGRPGNFPQSRPEPTAPPAEI
FGMGEEITSPPKQEQKDREQDPPLVSLKSLFGNDPLSQ$

Fig. 65B

3. 2003 CON A1 gag.OPT

ATGGGCGCCCGCCGCCTCCGTCCGTGCTGTCCGGCGGCAAGCTGGACGCCTGGGAGAAGATCCGCCTGCGCCTGCGCCCGGGCGGCAAGAAGAAGTACCGCCT
GAAGCACCTGGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCCTGAACCCCTCCCTGCTGGAGACCACCGAGGGCTGCCAGCAGATCATGG
AGCAGCTGCAGCCCGCCCTGAAGACCGGCACCGAGGAGCTGCGCTCCCTGTACAACACCGTGGCCACCCTGTACTGCGTGCACCAGCGCATC
GACGTGAAGGACACCAAGGAGGCCCTGGACAAGATCGAGGAGATCCAGAACAAGTCCAAGCAGAAGACCCAGCAGGCCGCCGCCGACACCGG
CAACTCCTCCAAGGTGTCCCAGAACTACCCCATCGTGCAGAACGCCCAGGGCCAGATGGTGCACCAGTCCCTGTCCCCGAGGACGCTGAACG
CCTGGGTGAAGGTGATCGAGGAGAAGGCCTTCTCCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCCCAGGACCTG
AACATGATGCTGAACATCGTGGGCGGCCATCCCCGCCGCCATGCCCAGATGCGCGAGCCCCGTGGGCGGCGACATCGCCGGCACCACCTCCAC
AGATCGGCTGGATGACCGGCAACCCGCCAACCCCGTGTCCATCCTGACATCTACACGCTGGATCATCCTGACCGGCTGGGCGACCCGCCCATG
TACTCCCCCGTGTCCATCCTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGAGAACGCCCACCCGGCCCACAAGCTGTGCGCCATG
GCCCGCGGCCACCCGGCCGCCACCCTGCAGGCCCTGCAGGAGATGATGACCGCGCGGGCGACGCCGCCCAGGCCGCCCATGTCC
CAGGTGCAGCACACCAACATCATGATGCAGCGCGGCAACTTCCGCGGCCAGAAGCGCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACCT
GGCCCGCAACTGCCGCGCGCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCA
ACTTCCTGGGCAAGATCTGGCCCTCCTCCAAGGGCCGCCCCGGCAACTTCCCCCAGTCCCGCCCCGAGCCCACCGCCCCGCCCGCCGAGATC
TTCGGCATGGGCGAGGAGATCACCTCCCCCCCCAAGCAGGAGCAGAAGGACCGCGAGCAGGACCCCCCCCTGGTGTCCCTGAAGTCCCTGTT
CGGCAACGACCCCCTGTCCCAGTAA

Fig. 65C

4. 2003_A1.anc gag.PEP

```
MGARASVLSGGKLDAWEKIRLRPGGKKKYRLKHLVWASRELERFALNPGLLETAEGCQQIMGQLQPALKTGTEELRSLYNTVATLYCVHQRI
EVKDTKEALDKIEEIQNKSKQKTQQAAADTGNSSKVSQNYPIVQNAQGQMVHQSLSPRTLNAWVKVIEEKAFSPEVIPMFSALSEGATPQDL
NMMLNIVGGHQAAMQMLKDTINEEAAEWDRLHPVHAGPIPPGQMREPRGSDIAGTTSTLQEQIGWMTGNPPIPVGDIYKRWIILGLNKIVRM
YSPVSILDIRQGPKEPFRDYVDRFFKTLRAEQATQEVKNWMTETLLVQNANPDCKSILRALGPGATLEEMTACQGVGGPGHKARVLAEAMS
QVQNTDIMMQRGNFRGPKRIKCFNCGKEGHLARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSSKGRPGNFPQSRPEPTAPPAEN
FGMGEEMISSPKQEQKDREQYPPLVSLKSLFGNDPLSQ$
```

Fig. 65D

2003_A1.anc gag.OPT

```
ATGGGCGCCCGCGCCTCCGTGCTGTCCGGCGGCAAGCTGGACGCGCCTGGGAGAAGATCCGCCTGCGCCCCGGCGGCAAGAAGAAGTACCGCCT
GAAGCACCTGGTGTGGGCCAGCCGCGAGCTGGAGCGCTTCGCCCTGAACCCCGGCCTGCTGGAGACCGCCGAGGGCTGCCAGCAGATCATGG
GCCAGCTGCAGCCCGCCCTGAAGACCGGCACCGAGGAGCTGCGCTCCCTGTACAACACCGTGGCCACCCTGTACTGCGTGCACCAGCGCATC
GAGGTGAAGGACACCAAGGAGGCCCTGGACAAGATCGAGGAGATCCAGAACAAGTCCAAGCAGAAGACCCAGCAGGCCGCCGACACCCGG
CAACTCCTCCTCCAAGGTGTCCCAGAACTACCCCATCGTGCAGAACGCCCAGGGCCAGATGGTGCACCAGTCCCTGTCCCCCCGCACCCTGAACG
CCTGGGTGAAGGTGATCGAGGAGAAGGCCTTCTCCCCCGAGGTGATCCCCATGTTCTCCGCCCTGAGCGAGGGCGCCACCCCCCAGGACCTG
AACATGATGCTGAACATCGTGGGCGGCCCCATCCCCCCGGCCAGATGCGCGAGCCCCGCGGCTCCGACATCGCCGGCACCACTTCCACCCTGCAGGAGC
AGATCGGCTGGATGACCGGCAACCCCCCCATCCCCGTGGGCGACATCTACAAGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATG
TACTCCCCCGTGTCCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCCTGCGCGCCGA
GCAGGCCACCCAGGAGGTGAAGAACTGGATGACCGAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGTCCATCCTGCGCGCCCTGG
GCCCCGGCGCCACCCTGGAGGAGATGACCGCGCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCAGCGTGCTGGCCGAGGCCATGTCC
CAGGTGCAGAACACCGACATCATGATGCAGCGCGGCAACTTCCGCGGCCCCAAGCGCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACCT
GGCCCGCAACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCA
ACTTCCTGGGCAAGATCTGGCCCTCCTCCAAGGGCCGCCCCGGCAACTTCCCCCAGAGCCGGCCCGAGCCCACCGCCCCCCCCGAGAAC
TTCGGCATGGGCGAGGAGATGATCTCCTCCCCCCAAGCAGGAGCAGAAGGACCGCGAGCAGTACCCCCCCCTGGTGTCCCTGAAGTCCCTGTTT
CGGCAACGACCCCCTGTCCCAGTAA
```

SUBSTITUTE SHEET (RULE 26)

Fig. 66A

5. 2003_CON_A2 gag.PEP

MGARASILSGGKLDAWEKIRLRPGGKKKYRLKHLVWASRELEKFSINPSLLETSEGCRQIIRQLQPALQTGTEELKSLYNTVAVLYCVHQRI
DVKDTKEALDKIEEEQNKCKQKTQHAAADTGNSSSSSQNYPIVQNAQGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFTALSEGATPQDL
NTMLNTVGGHQAAMQMLKDTINEEAAEWDRLHPVHAGPIPPGQMREPRGSDIAGTTSTLQEIGWMTSNPPIPVGEIYKRWIILGLNKIVRM
YSPVSILDIRQGPKEPFRDYVDRFFKTLRAEQATQEVKNWMTDTLLVQNANPDCKSILRALGPGATLEEMMTACQGVGGPSHKARVLAEAMS
QVQNTNIMMQRGNFRGQKRIKCFNCGKEGHLARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSNKGRPGNFPQSRTEPTAPPA
ENLRMGEEITSSLKQELKTREPYNPAISLKSLFGNDPLSQ$

Fig. 66B

2003_CON_A2 gag.OPT

ATGGGCGCCCGCGCCTCCATCCTGTCCGGCGGCAAGCTGGACGCGCCTGGGAGAAGATCCGCCTG
GAAGCACCTGGTGTGCGCCCGCCCGGCAAGAGAAGAAGTACCGCCT
GCCCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGAAGAGTTCTCCATCAACCCCTCCTTGTACAACACCGTGGCCGTGCTGTACTGCGTGCACCAGCGCATC
GACGTGAAGGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGAACAAGTGCAAGCAGAAGACCCAGCACGCCGCCGACACCGG
CAACTCCTCCTCCTCCAGAACTACCCCATCGTGCAGAACGCCCAGGGCCAGATGGTGCACCAGGCCATCTCCCCCCGCACCCTGAACG
CCTGGGTGAAGGTGGTGGAGGAGAAGGCCTTCTCCCCCGAGGTGATCCCCATGTTCACCGCCCTGTCCGAAGGCGCCACCCCCCAGGACCTG
AACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCT
GCACCCCGTGCACGCCGGCCCCATCCCCCCGGGCCAGATGCGCGAGCCCCGTGGGCCCCAAGAGCTCCGAGATCATCGCTGGACATCGTGCGCATG
TACTCCCCCGTGTCCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCCTGCGCGCCGA
GCAGGCCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGTCCATCCTGCGCGCCCTGG
GCCCCGGCGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCAACTTCCGGCCCAGGAGGCGCATCAAGTGTCC
CAGGTGCAGAACACCAACATCATGATGCAGCGCGGCAAGGGGGCTGCTGGAAGTGCGCCAGCCCGCAAGGGCCAACTTCCGCGGCCAGAAGCGCATCAAGTGCTTCAACTGCGGCAAGGAGGG
CCACCTGGCCCGCAACTGCCGCGCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGG
AGGCCAACTTCCTGGGCAAGATCTGGCCGAGAATCCCCGCAAGAAGGGCCGCCCCGGCAACTTCCCCCAGTCCCGCACCGAGCCCACCGCCCCCGCC
GAGAACCTGCGCATGGGCGAGGAGATCACCTCCTCCCTGAAGCAGGAGCTGAAGACCCGCGAGCCCTACAACCCCGCCATCTCCCTGAAGTC
CCTGTTCGGCAACGACCCCCTGTCCCAGTAA

Fig. 67A

6. 2003_CON_B gag.PEP

MGARASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERFAVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRI
EVKDTKEALEKIEEEQNKSKKKAQQAAADTGNSSQVSQNYPIVQNLQGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDL
NTMLNTVGGHQAAMQMLKETINEEAAEWDRLHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIVRM
YSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQNANPDCKTILKALGPAATLEEMTACQGVGGPGHKARVLAEAMS
QVTNSATIMMQRGNFRNQRKTVKCFNCGKEGHIAKNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSHKGRPGNFLQSRPEPTAPPE
ESFRFGEETTTPSQKQEPIDKELYPLAS$

Fig. 67B

2003 CON B gag.OPT

ATGGGCGCCCGCGCCTCCGTGCTGTCCGGCGGCGAGCTGGACCGCTGGGAGAAGATCCGCCTGCGCCCCGGCGGCAAGAAGAAGTACAAGCT
GAAGCACATCGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCGTGAACCCCGGCCTCCGAGGGCTGCCGAGGGGCTGCTTCGCCGTGCACCAGCGCATC
GCCAGCTGCAGCCCTCCCTGCAGACCGGCTCCGAGGAGCTGCGCTCCCTGTACAACACCGTGGCCACACTGTACTGCGTGCACCAGCGCATC
GAGGTGAAGGACACCAAGGAGGCCCTGGAGAAGATCGAGGAGGAGCAGAACAAGTCCAAGAAGAAGGCCCAGCAGGCCGCCGCCGACACCGG
CAACTCCTCCCAGGTGTCCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAGGCCATCTCCCCCCGCACCCTGAACG
CCTGGGTGAAGGTGGTGGAGGAGAAGGCCTTCTCCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCCCAGGACCTG
AACACCATGCTGAACACCGTGGGCGGCCACCAGGCGGCCATGCAGATGCTGAAGGAGACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCT
GCACCCCGTGCACGCCGGGATGACCAACAACCCCCCCATCGCCCCCGGGCCCCAGGGCCCCAGATGCGCGAGCCCCGTGGGGCGAGATCTACAAGCGCTGGATCATCCTGGGCCTG
AGATCGGCTGGATGACCAACAACCCCCCCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTACAAGACCCTGCGCGCCGA
GCAGGCCTCCCAGGAGGTGAAGAACTGGATGACCGAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCCCTGG
GCCCCGCCGCCACCCTGGAGGAGATGACCGCCTGCCAGGGCGTGGGCGGCCCAGGGCCACAAGGCCGTGAAGTGCTTCAACTGCGGCAAGGAGGG
CAGGTGACCAACTCCGCCACCATCATGATGCAGCGCGGCAACTTCCGCAACCAGCGCAAGACCGTGAAGTGCTTCAACTGCGGCAAGGAGGG
CCACATCGCCAAGAACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCC
AGGCCAACTTCCTGGGCAAGATCTGGCCCTCCCACAAGGGCCGCCCCGGCAACTTCCTGCAGTCCCGCCCCGAGCCCACCGCCCCCCCCGAG
GAGTCCTTCCGCTTCGGCGAGGAGACCACCACCCCCTCCCAGAAGCAGGAGCCCATCGACAAGGAGCTGTACCCCCTGGCCTCCTAA

Fig. 67C

7. 2003_B.anc gag.PEP

MGARASVLSGGKLDKWEKIRLRPGGKKKYKLKHIVWASRELERFAVNPGLLETSEGCRQILGQLQPALQTGSEELRSLYNTVATLYCVHQRI
EVKDTKEALDKIEEEQNKSKKKAQQAAADTGNSSQVSQNYPIVQNLQGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDL
NTMLNTVGGHQAAMQMLKETINEEAAEWDRLHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIVRM
YSPISILDIRQGPKEPFRDYVDRFYKTLRAEQASQDVKNWMTETLLVQNANPDCKTILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMS
QVTNSTTIMMQRGNFRDQRKIVKCFNCGKEGHIARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSHKGRPGNFLQSRPEPTAPPE
ESERFGEETTTPSQKQEPIDKELYPLASLKSLFGNDPSSQ$

Fig. 67D

2003_B.anc gag.OPT

ATGGGCGCCCGCGCCTCCGTGCTGTCCGGCGGCAAGCTGGACAAGTGGGAGAAGATCCGCCTGCGCCCCGGCGGCAAGAAGAAGTACAAGCT
GAAGCACATCGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCGTGAACCCCGGCCTGCTGGAGACCTCCGAGGGCTGCCGCCAGATCCTGG
GCCAGCTGCAGCCCGCCCTGCAGACCGGCTCCGAGGAGCTGCGCTCCCTGTACAACACCGTGGCCACCCTGTACTGCGTGCACCAGCGCATC
GAGGTGAAGGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGAACAAGTCCAAGAAGAAGGCCCAGCAGGCCGCCGCCGACACCGG
CAACTCCTCCCAGGTGTCCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAGGCCATCTCCCCCCGCACCCTGAACG
CCTGGGTGAAGGTGGTGGAGGAGAAGGCCTTCTCCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCCCAGGACCTG
AACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGAGACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCT
GCACCCCGTGCACGCCGGCCCCATCGCCCCCGGCCAGATGCGCGAGCCCCGTGGGGCCCCGAGATCTACAAGCGCTGGATCATCCTGGGCCTG
AGATCGGCCTGGATGAACACAACCCCCCATCCTGGACACTGGACAAGATCTCCCATCCTGGACACTGATGAAGAAGACTGGATGAGAGATGATGA
TACTCCCCCATCTCCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTACAAGACCCTGCGCGCCGA
GCAGGCCTCCCAGGACGTGAAGAACTGGATGACCGAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCCCTGG
GCCCCGCCGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCTGGCCACAAGGCTCGTGTGCTGGCCGAGGCCATGTCC
CAGGTGACCAACTCCACCACCATCATGATGCAGCGCGGCAACTTCCGCGACCAGCGCAAGATCGTGAAGTGCTTCAACTGCGGCAAGGAGGG
CCACATCGCCCGCAACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCC
AGGCCAACTTCCTGGGCAAGATCTGGCCCTCCCACAAGGGCCGCCCCGGCAACTTCCTGCAGTCCCGCCCCGAGCCCACCGCCCCCCCCGAG
GAGTCCTTCCGCGAGGAGACCACCACCCCCTCCCAGAAGCAGGAGCCCATCGACAAGGAGCTGTACCCCCTGGCCTCCCTGAAGTC
CCTGTTCGGCAACGACCCCTCCTCCCAGTAA

Fig. 68A

8. 2003_CON_C gag.PEP

MGARASILRGGKLDKWEKIRLRPGGKKHYMLKHLVWASRELERFALNPGLLETSEGCKQIIKQLQPALQTGTEELRSLYNTVATLYCVHEKI
EVRDTKEALDKIEEEQNKSQQKTQQAKAADGKVSQNYPIVQNLQGQMVHQAISPRTLNAWVKVIEEKAFSPEVIPMFTALSEGATPQDLNTM
LNTVGGHQAAMQMLKDTINEEAAEWDRLHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIAWMTSNPPIPVGDIYKRWIILGLNKIVRMYSP
VSILDIKQGPKEPFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQNANPDCKTILRALGPGATLEEMMTACQGVGGPSHKARVLAEAMSQAN
NTNIMMQRSNFKGPKRIVKCFNCGKEGHIARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSHKGRPGNFLQNRPEPTAPPAESFR
FEETTPAPKQEPKDREPLTSLKSLFGSDPLSQ$

Fig. 68B

2003_CON_C gag.OPT

ATGGGCGCCCGCGCCTCCATCCTGCGCGGCGGCAAGCTGGACAAGTGGGAGAAGATCCGCCTGCGCCCCGGCGGCAAGAAGCACTACATGCT
GAAGCACCTGGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCCTGAACCCCGGCCTGCTGGAGACCTCCGAGGGCTGCAAGCAGATCATCA
AGCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCTCCCTGTACAACACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATC
GAGGTGCGCGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAGACCCAGCAGGCCAAGGCCGCCGACGG
CAAGGTGTCCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAGGCCATCTCCCCGCGCACCCTGAACGCCTGGGTGA
AGGTGATCGAGGAGAAGGCCTTCTCCCCGGAGGTGATCCCCATGTTCACCGCCCTGTCCGAGGGCGCCACCCCCCAGGACCTGAACACCATG
CTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCTGCACCCCGT
GCACGCCGGCCCCATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCTCCGACATCGCCGGCACCACCTCCACCCTGCAGGAGCAGATCGCCT
GGATGACCTCCAACCCCCCCATCCCCGTGGGCGACATCTACAAGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCCCCC
GTGTCCATCCTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCCTGCGCGCCGAGCAGGCCAC
CCAGGACGTGAAGAACTGGATGACCGACACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCCCTGGGCCCCGGCG
CCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCTCCCACAAGGCCCGCGTGCTGGCCGAGGCCATGTCCCAGGCCAAC
AACACCAACATCATGATGCAGCGCTCCAACTTCAAGGGCCCCAAGCGCATCGTGAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCG
CAACTGCCGCGCCCCTCCCACAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCC
TGGGCAAGATCTGGCCCTCCCACAAGGGCCGCCCCGGCAACTTCCTGCAGAACCGCCCCGAGCCCACCGCCCCCCCCGCCGAGTCCTTCCGC
TTCGAGGAGACCACCCCCGCCCCCAAGCAGGAGCCCAAGGACCGCGAGCCCCTGACCTCCCTGAAGTCCCTGTTCGGCTCCGACCCCCTGTC
CCAGTAA

Fig. 68C

9. 2003 C.anc.gag.PEP

MGARASILRGGKLDTWEKIRLRPGGKKHYMIKHLVWASRELERFALNPGLLETSEGCKQIMKQLQPALQTGTEELRSLYNTVATLYCVHERI
EVRDTKEALDKIEEEQNKSQQKTQQAEAADGDNGKVSQNYPIVQNLQGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFTALSEGATPQDL
NTMLNTVGGHQAAMQMLKDTINEEAAEWDRLHPVHAGPVAPGQMREPRGSDIAGTTSTLQEQIAWMTSNPPIPVGDIYKRWIILGLNKIVRM
YSPVSILDIKQGPKEPFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQNANPDCKTILRALGPGATLEEMMTACQGVGGPGHKARVLAEAMS
QANNTNIMMQRSNFKGPKRIVKCFNCGKEGHIARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSHKGRPGNFLQSRPEPTAPPAE
SFREEETTPAPKQEPKDREPLTSLKSLFGSDPLSQ$

Fig. 68D

2003 C.anc.gag.OPT

ATGGGCGCCCGCGCCTCCATCCTGCGCGGCGGCAAGCTGGACACCTGGGAGAAGATCCGCCTGCGCCTGCGCCCCGGCGGCAAGAAGCACTACATGAT
CAAGCACCTGGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCCTGAACCCCGGCCTGCTGGAGACCTCCGAGGGCTGCAAGCAGATCATGA
AGCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCTCCCTGTACAACACCGTGGCCACCCTGTACTGCGTGCACGAGCGCATC
GAGGTGCGCGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAGACCCAGCAGGCCGAGGCCGCCGACGG
CGACAACGGCAAGGTGTCCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAGGCCATCTCCCCCCGCACCCTGAACG
CCTGGGTGAAGGTGGTGGAGGAGAAGGCCTTCTCCCCCGAGGTGATCCCCATGTTCACCGCCCTGTCCGAGGGCGCCACCCCCCAGGACCTG
AACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCT
GCACCCCGTGCACGCCGGCCCCGTGGCCCCCGGCCAGATGCGCGAGCCCCGCGGCTCCGACATCGCCGGCACCACCTCCACCCTGCAGGAGC
AGATCGCCTGGATGACCTCCAACCCCCCCATCCCCGTGGGCGACATCTACAAGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATG
TACTCCCCCGTGTCCATCCTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCCTGCGCGCCGA
GCAGGCCACCCAGGACGTGAAGAACTGGATGACCGACACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCCCTGG
GCCCCGGCGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTGCTGGCCGAGGCCATGTCC
CAGGCCAACAACACCAACATCATGATGCAGCGCTCCAACTTCAAGGGCCCCAAGAGGATCGTGAAGTGCTTCAACTGCGGCAAGGAGGGCCA
CATCGCCCGCAACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGG
CCAACTTCCTGGGCAAGATCTGGCCCTCCCACAAGGGCCGCCCCGGCAACTTCCTGCAGTCCCGCCCCGAGCCCACCGCCCCCCCCGCCGAG
TCCTTCCGCGAGGAGGAGACCACCCCCGCCCCCAAGCAGGAGCCCAAGGACCGCGAGCCCCTGACCTCCCTGAAGTCCCTGTTCGGCTCCGA
CCCCCTGTCCCAGTAA

Fig. 69A

10. 2003_CON_D gag.PEP

MGARASVLSGGKLDAWEKIRLRPGGKKKYRLKHIVWASRELERFALNPGLLETSEGCKQIIGQLQPAIQTGSEELRSLYNTVATLYCVHERI
EVKDTKEALEKIEEQNKSKKKAQQAAADTGNSSQVSQNYPIVQNLQGQMVHQAISPRTLNAWVKVIEEKAFSPEVIPMFSALSEGATPQDL
NTMLNTVGGHQAAMQMLKETINEEAAEWDRLHPVHAGPVAPGQMREPRGSDIAGTTSTLQEIGWMTSNPPIPVGEIYKRWIILGLNKIVRM
YSPVSILDIRQGPKEPFRDYVDRFYKTLRAEQASQDVKNWMTETLLVQNANPDCKTILKALGPEATLEEMTACQGVGGPSHKARVLAEAMS
QATNSAAVMMQRGNFKGPRKIIKCFNCGKEGHIAKNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSHKGRPGNFLQSRPEPTAPPA
ESFGFGEEITPSQKQEQKDKELYPLTSLKSLFGNDPLSQS

Fig. 69B

2003_CON_D gag.OPT

ATGGGCGCCCGCGCCTCCGTGCTGTCCGGCGGCAAGCTGGACGCCTGGGAGAAGATCCGCCTGCGCCCCGGCGGCAAGAAGAAGTACCGCCT
GAAGCACATCGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCCTGAACCCCGGCCTCCTGGAGACCTCGGAGGGCTGCAAGCAGATCATCG
GCCAGCTGCAGCCCGCCATCCAGACCGGCTCCGAGGAGCTGCGCTCCCTGTACAACACCGTGGCCACCCTGTACTGCGTGCACGAGCGCATC
GAGGTGAAGGACACCAAGGAGGCCCTGGAGAAGATCGAGGAGCAGAACAAGTCCAAGAAGAAGGCCCAGCAGGCCGCCGCCGACACCGGG
CAACTCCTCCCAGGTGTCCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCCCCTGTCCCCCCGAGGTGATCATGCCAACG
CCTGGGTGAAGGTGATCGAGGAGAAGGCCTTCTCCCCCGAGGTGATCCCCATGCGCGCGAGGAGCCATCAACGAGAGGCCGCCGCCGGGGGACCGCCT
AACACCATGCTGAACACCGTGGGCGGCCACCCGGCGCCGCCGTGGCCCCCGGCGGCCAGATGCGCGAGCCCCGCGGCTCCGACGCCGCCCGCCAGGAGC
AGATCGGCTGGATGACCTCCAATCCTGGACATCCGCCAGGGCCCCAGAGACCCGGCCGCCTGCAGGGCGCGGCCCGAGGCCCTGG
GCCCCGAGGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCTCCCACAAGGCCCGCGTGCTGGCCGAGGCCATGTCC
CAGGCCACCAACTCCGCCGCCGTGATGATGCAGCGCGGCAACTTCAAGGGCCCGCGCAAGATCATCAAGTGCTTCAACTGCGGCAAGGAGGG
CCACATCGCCAAGAACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCC
AGGCCAACTTCCTGGGCAAGATCTGGCCCTCCCACAAGGGCCGCCCCGGCAACTTCCTGCAGTCCCGCCCCGAGCCCACCGCCCCCCCCGCC
GAGTCCTTCGGCTTCGGCGAGGAGATCACCCCCTCCCAGAAGCAGGAGCAGAAGGACAAGGAGCTGTACCCCCTGACCTCCCTGAAGTCCCT
GTTCGGCAACGACCCCCTGTCCCAGTAA

Fig. 70A

11. 2003_CON_F gag.PEP

MGARASVLSGGKLDAWEKIRLRPGGKKKYRMKHLVWASRELERFALDPGLLETSEGCQKIIGQLQPSLQTGSEELRSLYNTVAVLYCVHQKV
EVKDTKEALEKLEEEQNKSQQKTQQAAADKGVSQNYPIVQNLQGQMVHQAISPRTLNAWVKVIEEKAFSPEVIPMFSALSEGATPQDLNTML
NTVGGHQAAMQMLKDTINEEAAEWDRLHPVHAGPIPPGQMREPRGSDIAGTTSTLQEQIQWMTSNPPVPVGDIYKRWIILGLNKIVRMYSPV
SILDIRQGPKEPFRDYVDRFFKTLRAEQATQEVKGWMTDTLLVQNANPDCKTILKALGPGATLEEMMTACQGVGGPGHKARVLAEAMSQATN
TAIMMQKSNFKGQRRIVKCFNCGKEGHIAKNCRAPRKKGCWKCGREGHQMKDCTERQANFLGKIWPSNKGRPGNFLQSRPEPTAPPAESFGF
REEITPSPKQEQKDEGLYPPLASLKSLFGNDP$

Fig. 70B

2003_CON_F gag.OPT

ATGGGCGCCCGCGCCGTCCTGTCCGGCGGCAAGCTGGACGCCTGGGAGAAGATCCGCCTGCGCCCCGGCGGCAAGAAGAAGTACCGCAT
GAAGCACCTGGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCCTGGACCCCGGCCTCCTGGAGACCTCCGAGGGCTGCCAGAAGATCATCG
GCCAGCTGCAGCCCTCCCTGCAGACCGGCTCCGAGGAGCTGCGCTCCCTGTACAACACCGTGGCCGTGCTGTACTGCGTGCACCAGAAGGTG
GAGGTGAAGGACACCAAGGAGGCCCTGGAGAAGCTGGAGGAGGAGCAGAACAAGTCCCAGCAGAAGACCCAGCAGGCCGCCGACAAGGG
CGTGTCCCAGAACTACCCCATCGTGCAGAACCTGCAGGGTCAGATGGTGCACCAGGCCATCTCCCCGCGCACCCTGAACGCCTGGGTGAAGG
TGATCGAGGAGAAGGCCTTCTCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCCCAGGACCTGAACACCATGCTG
AACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCTGCACCCCGTGCA
CGCCGGCCCCATCCCCCCGGGCCAGATGCGCGAGCCCCGCGGCTCCGACATCGCCGGCACCACCTCCACCCTGCAGGAGCAGATCCAGTGGA
TGACCTCCAACCCCCCGGTGCCCGTGGGCGACATCTACAAGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGCGAGATGTACTCCCCGGTG
TCCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCCTGCGCGCCGAGCAGGCCACCCA
GGAGGTGAAGGGCTGGATGACCGACACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCCCTGGGCCCCGGCGCCA
CCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCATCGTGCTGCCAGGCCATGCGCGAGGAGGGCCAAGGAGAGGGCCAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAGAA
CTGCCGCGCCCCCGCCTCCAACAAGGGCTGCTGGAAGTGCGGCCGCGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTGG
GCAAGATCTGGCCCTCCAACAAGGGCCGCCCCGGCAACTTCCTGCAGTCCCGCCCCGAGCCCACCGCCCCGCCCGCCGAGTCCTTCGGCTTC
CGCGAGGAGATCACCCCCTCCCCCAAGCAGGAGCAGAAGGACGAGGGCCTGTACCCCCCCCTGGCCTCCCTGAAGTCCCTGTTCGGCAACGA
CCCCTAA

SUBSTITUTE SHEET (RULE 26)

Fig. 71A

12. 2003 CON G gag.PEP

MGARASVLSGGKLDAWEKIRLRPGGKKKYRMKHLVWASRELERFALNPDLLETAEGCQQIMGQLQPALQTGTEELRSLFNTVATLYCVHQRI
EVKDTKEALEEVEKIQKKSQQKTQQAAMDEGNSSQVSQNYPIVQNAQGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDL
NTMLNTVGGHQAAMQMLKDTINEEAAEWDRMHPQQAGPIPPGQIREPRGSDIAGTTSTLQEQIRWMTSNPPIPVGEIYKRWIILGLNKIVRM
YSPVSILDIRQGPKEPFRDYVDRFFKTLRAEQATQEVKGWMTDTLLVQNANPDCKTILRALGPGATLEEMMTACQGVGGPSHKARVLAEAMS
QASGAAAIMMQKSNFKGPRRTIKCFNCGKEGHLARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSNKGRPGNFLQNRPEPTAPP
AESFGFGEEIAPSPKQEQKEKELYPLASLKSLFGSDP$

Fig.71B

2003 CON G gag.OPT

ATGGGCGCCCGCGCCTCCGTGCTGTCCGGCGGCAAGCTGGACGCCTGGGAGAAGATCCGCCTGCGCCCCGGCGGCAAGAAGAAGTACCGCAT
GAAGCACCTGGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCCTGAACCCCGACCTGCTGGAGACCGCCGAGGGCTGCCAGCAGATCATGG
GCCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCTCCCTGTTCAACACCGTGGCCACCCTGTACTGCGTGCACCAGCGCATC
GAGGTGAAGGACACCAAGGAGGCCCTGGAGGAGGTGGAGAAGATCCAGAAGAAGTCCCAGCAGAAGACCCAGCAGGCCGCCATGGACGAGGG
CAACTCCTCCCAGGTGTCCCAGAACTACCCCATCGTGCAGAACGCCCAGGGCCAGATGGTGCACCAGGCCATCTCCCCGCGCACCCTGAACG
CCTGGGTGAAGGTGGTGGAGGAGAAGGCCTTCTCCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCCCAGGACCTG
AACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCAT
GCACCCCCAGCAGGCCGGCCCCATCCCCCCCGGCCAGATCCGCGAGCCCCGTGGGCTCCGACATCGCCGGCACCACCTCCACCCTGCAGGAGC
AGATCCGCTGGATGACCTCCAACCCCCCCATCCCCGTGGGCGAGATCTACAAGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATG
TACTCCCCCGTGTCCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCCTGCGCGCCGA
GCAGGCCACCCAGGAGGTGAAGGGCTGGATGACCGACACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCCCTGG
GCCCCGGCGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCTCCCACAAGGCCCGCGTGCTGGCCGAGGCCATGTCC
CAGGCCTCCGGCGCCGCCGCCATCATGATGCAGAAGTCCAACTTCAAGGGCCCCCGCCGCACCATCAAGTGCTTCAACTGCGGCAAGGA
GGGCCACCTGGCCCGCAACTGCCGCGCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGC
GCCAGGCCAACTTCCTGGGCAAGATCTGGCCCTCCAACAAGGGCCGCCCCGGCAACTTCCTGCAGAACCGCCCCGAGCCCACCGCCCCCCCC
GCCGAGTCCTTCGGCTTCGGCGAGGAGATCGCCCCCTCCCCCAAGCAGGAGCAGAAGGAGAAGGAGCTGTACCCCCTGGCCTCCCTGAAGTC
CCTGTTCGGCTCCGACCCCTAA

Fig. 72A

13. 2003 CON H gag.PEP

MGARASVLSGGKLDAWEKIRLRLRPGGKKKYRLKHLVWASRELERFALNPGLLETAEGCLQIIEQLQPAIKTGTEELQSLFNTVAVLYCVHQRI
DVKDTKEALGKIEEIQNKSQQKTQQAAADKEKDNKVSQNYPIVQNAQGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDL
NAMLNTVGGHQAAMQMLKDTINEEAAEWDRLHPVHAGPIPPGQMREPRGSDIAGTTSTLQEQIAWMTGNPPIPVGDIYKRWIILGLNKIVRM
YSPVSILDIKQGPKEPFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQNANPDCKTILRALGQGASIEEMMTACQGVGGPSHKARVLAEAMS
QVTNANAAIMMQKGNFKGPRKIVKCFNCGKEGHIARNCRAPRKKGCWKCGREGHQMKDCTERQANFLGKIWPSSKGRPGNFLQSRPEPTAPP
AESFGFGEEMTPSPKQELKDKEPPLASLRSLFGNDPLSQS

Fig. 72B

2003 CON H gag.OPT

ATGGGCGCCCGCGCCTCCGTGCTGTCCGGCGGCAAGCTGGACGCCTGGGAGAAGATCCGCCTGCGCCGGCAAGAAGAAGTACCGCCT
GAAGCACCTGGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCCTGAACCCCGGCCTGCTGGAGACCGCCGAGGGCTGCCTGCAGATCATCG
AGCAGCTGCAGCCCGCCATCAAGACCGGCACCGAGGAGCTGCAGTCCCTGTTCAACACCGTGGCCGTGCTGTACTGCGTGCACCAGCGCATC
GACGTGAAGGACACCAAGGAGGCCCTGGGCAAGATCGAGGAGATCCAGAACAAGTCCCAGCAGAAGCAGCAGGCCGCCGCCGACAAGGA
GAAGGACAACAAGGTGTCCCAGAACTACCCCATCGTGCAGAACGCCCAGGGCCAGATGGTGCACCAGGCCATCTCCCCCCGCACCCTGAACG
CCTGGGTGAAGGTGGTGGAGGAGAAGGCCTTCTCCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCCCAGGACCTG
AACGCCATGCTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCT
GCACCCCGTGCACGCCGGCCCCATCCCCCCCGGCCAGATGCGCGAGCCCCGTGGTCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATG
TACTCCCCCGTGTCCATCCTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCCTGCGCGCCGA
GCAGGCCACCCAGGACGTGAAGAACTGGATGACCGACACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCCCTGG
GCCAGGGCGCCTCCATCGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTGCTGGCCGAGGCCATGTCC
CAGGTGACCAACGCCAACGCCGCCATCATGATGCAGAAGGGCAACTTCAAGGGCCCCCGCAAGATCGTGAAGTGCTTCAACTGCGGCAAGGA
GGGCCACATCGCCCGCAACTGCCGCGCCCCTCGCAAGAAGGGCTGCTGGAAGTGCGGCCGCGAGGGCCACCAGATGAAGGACTGCACCGAGC
GCCAGGCCAACTTCCTGGGCAAGATCTGGCCCTCCTCCAAGGGCCGCCCCGGCAACTTCCTGCAGTCCCGCCCCGAGCCCACCGCCCCCCCC
GTTCGGCCAACGACCCCCTGTCCCAGTAA

Fig. 73A

14. 2003 CON_K gag.PEP

MGARASVLSGGKLDTWEKIRLRPGGKKKYRLKHLVWASRELERFALNPSLLETTEGCRQIIRQLQPSLQTGSEELKSLFNTVATLYCVHQRI
EVRDTKEALDKLEEEQNKSQQKTQQETADKGVSQNYPIVQNLQGQMVHQALSPRTLNAWVKVIEEKAFSPEVIPMFSALSEGATPQDLNTML
NTVGGHQAAMQMLKDTINEEAAEWDRLHPVHAGPIPPGQMREPRGSDIAGTTSTLQEQITWMTSNPPVPVGEIYKRWIILGLNKIVRMYSPV
SILDIRQGPKEPERDYVDRFFKTLRAEQATQEVKNWMTDTLLVQNANPDCKTILKALGPGASLEEMTACQGVGGPGHKARILAEAMSQVTN
TAVMMQRGNFKGQRKIIKCFNCGKEGHIARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSNKGRPGNFLQSRPEPTAPPAESFGF
GEEITPSPRQETKDKEQGPPLTSLKSLFGNDPLSQ$

Fig. 73B

2003 CON_K gag.OPT

ATGGGCGCCCGCGCCTCCGTGCTGTCCGGCGGCAAGCTGGACACCTGGGAGAAGATCCGCCTGCGCCCCGGCGGCAAGAAGAAGTACCGCCT
GAAGCACCTGGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCCTGAACCCCTCCCTGCTGGAGACCACCGAGGGCTGCCGCCAGATCATCC
GCCAGCTGCAGCCCTCCCTGCAGACCGGCTCCGAGGAGCTGAAGTCCCTGTTCAACACCGTGGCCACCCTGTACTGCGTGCACCAGCGCATC
GAGGTGCGCGACACCAAGGAGGCCCTGGACAAGCTGGAGGAGGAGCAGAACAAGTCCCAGCAGAAGACCCAGCAGGAGACCGCCGACAAGGG
CGTGTCCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAGGCCCTGTCCCCCCGCACCCTGAACGCCTGGGTGAAGG
TGATCGAGGAGAAGGCCTTCTCCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCCCAGGACCTGAACACCATGCTG
AACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCTGCACCCCGTGCA
CGCCGGCCCCATCCCCCCGGCCAGATGCGCGAGCCCCGCGGCTCCGACATCGCCGGCACCACCTCCACCCTGCAGGAGCAGATCACCTGGA
TGACCTCCAACCCCCCCGTGCCCGTGGGCGAGATCTACAAGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCCCCCGTG
TCCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCGAGCGCGACTACGTGGACCGCTTCTTCAAGACCCTGCGCGCCGAGCAGGCCACCCA
GGAGGTGAAGAACTGGATGACCGACACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCCCTGGGCCCCGGCGCCT
CCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCAGGCCACAAGGCCCGCATCCTGGCCGAGGCCATGTCCCAGGTGACCAAC
ACCGCCGTGATGATGCAGCGCGGCAACTTCAAGGGCCAGCGCAAGATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAA
CTGCCGCGCCCCGCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTGG
GCAAGATCTGGCCCTCCAACAAGGGCCGCCCCGGCAACTTCCTGCAGTCCCGCCCAGAGTCCTTCGGCTTCGGCGAGGAGATCACCCCCTCC
CCCCGCCAGGAGACCAAGGACAAGGAGCAGGGCCCCCCCCTGACCTCCCTGAAGTCCCTGTTCGGCAACGA
CCCCCTGTCCCAGTAA

Fig. 74A

15. 2003_CON_01_AE gag.PEP

MGARASVLSGGKLDAWEKIRLRPGGKKKYRMKHLVVASRELERFALNPGLLETAEGCQQIIEQLQSTLKTGSEELKSLFNTVATLWCVHQRI
EVKDTKEALDKIEEVQNKSQQKTQQAAAGTGSSSKVSQNYPIVQNAQGQMVHQPLSPRTLNAWVKVVEEKGFNPEVIPMFSALSEGATPQDL
NMMLNIVGGHQAAMQMLKETINEEAAEWDRVHPVHAGPIPPGQMREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGDIYKRWIILGLNKIVRM
YSPVSILDIRQGPKEPFRDYVDRFYKTLRAEQATQEVKNWMTETLLVQNANPDCKSILKALGTGATLEEMTACQGVGGPSHKARVLAEAMS
QAQHANIMMQRGNFKGQKRIKCFNCGKEGHLARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSNKGRPGNFPQSRPEPTAPPAEN
WGMGEEITSLPKQEQKDKEHPPPLVSLKSLFGNDPLSQ$

Fig. 74B

2003_CON_01_AE gag.OPT

ATGGGCGCCCGCGCCAGCGTGCTGTCCGGCGGCAAGCTGGACGCCTGGGAGAAGATCCGCCTGCGCCCCGGCGGCAAGAAGAAGTACCGCAT
GAAGCACCTGGTGGTGGCCTCCAGGGAGCTGGAGCGCTTCGCCCTGAACCCCGGCCTGCTGGAGACCGCCGAGGGCTGCCAGCAGATCATCG
AGCAGCTGCAGTCCACCCTGAAGACCGGCTCCGAGGAGCTGAAGTCCCTGTTCAACACCGTGGCCACCCTGTGGTGCGTGCACCAGCGCATC
GAGGTGAAGGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGTGCAGAACAAGTCCCAGCAGAAGACCCAGCAGGCCGCCGCCGGCACCGG
CTCCTCCTCCAAGGTGTCCCAGAACTACCCCATCGTGCAGAACGCCCAGGGCCAGATGGTGCACCCCCTGTCCCCGCGGGACCCTGAACG
CCTGGGTGAAGGTGGTGGAGGAGAAGGGCTTCAACCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCCCAGGACCTG
AACATGATGCTGAACATCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGAGACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGT
GCACCCCGTGCACGCCGGCCCCATCCCCCCCGGCCAGATGCGCGAGCCCCGCGGCTCCGACATCGCCGGCACCACCTCCACCCTGCAGGAGC
AGATCGGCTGGATGACCAACAACCCCCCCATCCCCGTGGGCGACATCTACAAGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATG
TACTCCCCCGTGTCCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTACAAGACCCTGCGCGCCGA
GCAGGCCACCCAGGAGGTGAAGAACTGGATGACCGAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGTCCATCCTGAAGGCCCTGG
GCACCGGCGCCACCCTGGAGGAGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTGCTGGCCGAGGCCATGTCC
CAGGCCCAGCACGCCAACATCATGATGCAGCGCGGCAACTTCAAGGGCCAGAAGCGCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACCT
GGCCCGCAACTGCCGCGCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCA
ACTTCCTGGGCAAGATCTGGCCCTCCAACAAGGGCCGCCCCGGCAACTTCCCCCAGTCCCGGCCCGAGCCCACCGCCCCCGCCGGAGAAC
TGGGGCATGGGCGAGGAGATCACCTCCCTGCCCAAGCAGGAGCAGAAGGACAAGGAGCACCCCCCCCTGGTGTCCCTGAAGTCCCTGTT
CGGCAACGACCCCCTGTCCCAGTAA

SUBSTITUTE SHEET (RULE 26)

Fig. 75A

16. 2003_CON_02_AG gag.PEP

MGARASVLSGGKLDAWEKIRLRPGGKKKYRLKHLVWASRELERFALNPGLLETAEGCQQIMEQLQSALRTGSEELKSLYNTVATLWCVHQRI
DIKDTKEALDKIEEVQNKSKQKTQQAAAATGSSSQNYPIVQNAQGQMTHQSMSPRTLNAWVKVIEEKAFSPEVIPMFSALSEGATPQDLNMM
LNIVGGHQAAMQMLKDTINEEAAEWDRVHPVHAGPIPPGQMREPRGSDIAGTTSTLQEQIGWMTSNPPIPVGEIYKRWIVLGLNKIVRMYSP
VSILDIRQGPKEPFRDYVDRFFKTLRAEQATQEVKNWMTETLLVQNANPDCKSILRALGPGATLEEMTACQGVGGPGHKARVLAEAMSQVQ
QSNIMMQRGNFRGQRTIKCFNCGKEGHLARNCKAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSSKGRPGNFPQSRPEPTAPPAESFGM
GEEITSSPKQEPRDKGLYPPLTSLKSLFGNDP$

Fig. 75B

2003_CON_02_AG gag.OPT

ATGGGCGCCCGCGCCTCCGTGCTGTCCGGCGGCAAGCTGGACGCGCCTGGGAGAAGATCCGCCTGCGCCCGGCGGCAAGAAGAAGTACCGCCT
GAAGCACCTGGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCCTGAACCCCGGCCTGCTGGAGACCCGCCGCGCCGAGGGCTGCCAGCAGATCATGG
AGCAGCTGCAGTCCGCCCTGCGCACCGGCTCCGAGGAGCTGAAGTCCCTGTACAACACCGTGGCCACCCTGTGGTGCGTGCACCAGCGCATC
GACATCAAGGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGTGCAGAACAAGTCCAAGCAGAAGACCCAGCAGGCCGCCGCCGCCACCGG
CTCCTCCTCCCAGAACTACCCCATCGTGCAGAACGCCCAGGGCCAGATGACCCACCAGTCCATGTCCCCGAGGGCCCACCCTGAACGCCTGGGTGA
AGGTGATCGAGGAGAAGGCCTTCTCCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCCCAGGACCTGAACATGATG
CTGAACATCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCCGT
GCACGCCGGCCCCATCCCCCCCGGCCAGATGCGCGAGCCCCGCGGCTCCGACATCGCTGGCGGCACCACCTCCACCCTGCAGGAGCAGATCGGCT
GGATGACCTCCAACCCCCCCATCCCCGTGGGCGAGATCTACAAGCGCTGGATCGTGCTGGGCCTGAACAAGATCGTGCGCATGTACTCCCCC
GTGTCCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCCTGCGCGCCGAGCAGGCCAC
CCAGGAGGTGAAGAACTGGATGACCGAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGTCCATCCTGCGCGCCCTGGGCCCCGGCG
CAGTCCAACATCATGATGCAGCGCGGCAACTTCCGCGGCCAGCGCACCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACCTGGCCCGCAA
CTGCAAGGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTGG
GGCAAGATCTGGCCCTCCTCCAAGGGCCGCCCCGGCAACTTCCCCCAGTCCCGCCCCGAGCCCACCGCCCCGCCAGTCCTTCGGCATG
GGCGAGGAGATCACCTCCTCCCCCAAGCAGGAGCCCCGCGACAAGGGCCTGTACCCCCCCCTGACCTCCCTGAAGTCCCTGTTCGGCAACGA
CCCCTAA

Fig. 76A

17. 2003_CON_03_ABG_gag.PEP

MGARASVLSGGKLDAWEKIRLRPGGKKKYRIKHLVWASRELERFALNPSLLETSEGCQQILEQLQPTLKTGSEELKSLYNTVATLYCVHQRI
EIKDTKEALDKIEEIQNKSKQKTQQAATGTGSSSKVSQNYPIVQNAQGQMTHQSMSPRTLNAWVKVIEEKAFSPEVIPMFSALSEGATPQDL
NMMLNIVGGHQAAMQMLKDTINEEAAEWDRLHPAQAGPFPPGQMREPRGSDIAGTTSTLQEQIGWMTSNPPIPVGDIYKRWIILGLNKIVRM
YSPVSILDIRQGPKEPFRDYVDRFFKTLRAEQATQDVKNWMTETLLVQNANPDCKTILRALGSGATLEEMMTACQGVGGPGHKARVLAEAMS
QVQNANIMMQKSNFRGPKRIKCFNCGKDGHLARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGRIWPSSKGRPGNFPQSRPEPSAPPAEN
FGMGEEITPSLKQEQKDREQHPPSISLKSLFGNDPLSQ$

Fig. 76B

2003_CON_03_ABG_gag.OPT

ATGGGCGCCCGCGCCTCCGTGCTGTCCGGCGGCAAGCTGGACGCCTGGGAGAAGATCCGCCTGCGCCCCGGCGGCAAGAAGAAGTACCGCAT
CAAGCACCTGGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCCTGAACCCCTCCCTGCTGGAGACCTCCGAGGGCTGCCAGCAGATCCTGG
AGCAGCTGCAGCCCACCCTGAAGACCGGCTCCGAGGAGCTGAAGTCCCTGTACAACACCGTGGCCACCCTGTACTGCGTGCACCAGCGCATC
GAGATCAAGGACACCAAGGAGGCCCTGGACAAGATCGAGGAGATCCAGAACAAGTCCAAGCAGAAGACCCAGCAGGCCGCCACCGGCACCGG
CTCCTCCTCCAAGGTGTCCCAGAACTACCCCATCGTGCAGAACGCCCAGGGCCAGATGACCCACCAGTCCATGTCCCCGCGCACCCTGAACG
CCTGGGTGAAGGTGATCGAGGAGAAGGCCTTCTCCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCCCAGGACCTG
AACATGATGCTGAACATCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCT
GCACCCCGCCCAGGCCGGCCCCTTCCCCCCAGGCCAGATGCGCGAGCCCCGTGGCTCCGACATCGCCGGCACCACCTCCACCCTGCAGGAGC
AGATCGGCTGGATGACCTCCAACCCCCCCATCCCCGTGGGCGACATCTACAAGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATG
TACTCCCCCGTGTCCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCCTGCGCGCCGA
GCAGGCCACCCAGGACGTGAAGAACTGGATGACCGAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCCCTGG
GCTCCGGCGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCATCAAGTGCTTCAACTGCGGCAAGGACGGCCACCT
GGCCCGCAACTGCCGCGCCCCGCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCA
ACTTCCTGGGCCGCATCTGGCCCTCCTCCAAGGGCCGCCCCGGCAACTTCCCCCAGTCCCGCCCCGAGCCCTCCGCCCCCCCGCCGAGAAC
TTCGGCATGGGCGAGGAGATCACCCCCTCCCTGAAGCAGGAGCAGAAGGACCGCGAGCAGCACCCGCCCTCCATCTCCCTGAAGTCCCTGTT
CGGCAACGACCCCCTGTCCCAGTAA

Fig. 77A

18. 2003_CON_04_CFX gag.PEP

```
MGARASVLSGGKLDAWERIRLRPGGKKKYRLKHLVWASRELERFALNPGLLETAEGCQQLMEQLQSTLKTGSEELKSLFNTIATLWCVHQRI
DVKDTKEALDKVEEMQNKSKQKTQQAAADTGGSSNVSQNYPIVQNAQGQMVHQSISPRTLNAWVKVIEEKAFSPEVIPMFSALSEGATPQDL
NMMLNIVGGHQAAMQMLKDTINEEAAEWDRAHPVHAGPIPPGQMREPRGSDIAGTTSTLQEQIGWMTSNPPIPVGEIYKRWIILGLNKIVRM
YSPVSILDIRQGPKEPFRDYVDRFFKCLRAEQATQEVKNWMTETLLVQNANPDCKSILKALGTGATLEEMMTACQGVGGPSHKARVLAEAMS
QASNAAAAIMMQKSNFKGQRRIIKCFNCGKEGHLARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGRMWPSSKGRPGNFLQSRPEPTAPP
AESLEMKEETTSSPKQEPRDKELYPLTSLKSLFGSDPLSQS
```

Fig. 77B

2003_CON_04_CFX gag.OPT

```
ATGGGCGCCCGCGCCTCCGTGCTGTCCGGCGGCAAGCTGGACGCCTGGGAGCGCATCCGCCTGCGCCCCGGCGGCAAGAAGAAGTACCGCCT
GAAGCACCTGGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCCTGAACCCCGGCCTGCTGGAGACCGCCGAGGGCTGCCAGCAGCTGATGG
AGCAGCTGCAGTCCACCCTGAAGACCGGCTCCGAGGAGCTGAAGTCCCTGTTCAACACCATCGCCACCCTGTGGTGCGTGCACCAGCGCATC
GACGTGAAGGACACCAAGGAGGCCCTGGACAAGGTGGAGGAGATGCAGAACAAGTCCAAGCAGAAGACCCAGCAGGCCGCCGCCGACACCGG
CGGCTCCTCCAACGTGTCCCAGAACTACCCCATCGTGCAGAACGCCCAGGGCCAGATGGTGCACCAGTCCATCTCCCCCCGCACCCTGAACG
CCTGGGTGAAGGTGATCGAGGAGAAGGCCTTCTCCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCCCAGGACCTG
AACATGATGCTGAACATCGTCGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGC
CCACCCCGTGCACGCCGGCCCCATCCCCCCAGGCCAGATGCGCGAGCCCCGCGGCTCCGACATCGCCGGCACCACCTCCACCCTGCAGGAGC
AGATCGGCTGGATGACCTCCAATCCCCCGATCCCCGTGGGCGAGATCTACAAGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGCGCCA
TACTCCCCCGTGTCCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGTGCCTGCGCGCCGA
GCAGGCCACCCAGGAGGTGAAGAACTGGATGACCGAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGTCCATCCTGAAGGCCCTGG
GCACCGGCGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCTCCCACAAGGCCCGCGTGCTGGCCGAGGCCATGTCC
CAGGCCTCCAACGCCGCCGCCGCCATCATGATGCAGAAGTCCAACTTCAAGGGCCAGCGCCGCATCATCAAGTGCTTCAACTGCGGCAAGGA
GGGCCACCTGGCCCGCAACTGCCGCGCCCCGCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGC
GCCAGGCCAACTTCCTGGGCCGCATGTGGCCCTCCTCCAAGGGCCGCCCAGGCAACTTCCTGCAGTCCCGCCCAGAGCCCACCGCCCCCCCC
GCCGAGTCCCTGGAGATGAAGGAGGAGACCACCTCCTCCCCCAAGCAGGAGCCCCGCGACAAGGAGCTGTACCCCCTGACCTCCCTGAAGTC
CCTGTTCGGCTCCGACCCCCTGTCCCAGTAA
```

Fig. 78A

19. 2003_CON_06_CPX_gag.PEP

MGARASVLSGGKLDEWEKIRLRPGGKKKYRLKHLVWASRELERFALNPGLLETAEGCQQIIEQLQSALKTGSEELKSLYNTVATLYCVHQRI
KVTDTKEALDKIEEIQNKSKQKAQQAAAATGNSSNLSQNYPIVQNAQGQMVHQAISPRTLNAWVKVIEEKAFSPEVIPMFSALSEGATPQDL
NMMLNIVGGHQAAMQMLKDTINEEAAEWDRVHPVHAGPIPPGQMREPRGSDIAGTTSTLQEQIGWMTSNPPIPVGEIYKRWIILGLNKIVRM
YSPVSILDIRQGPKEPFRDYVDRFFKTLRAEQATQEVKNWMTDTLLVQNANPDCKTILKALGPGATLEEMMTACQGVGGPGHKARVLAEAMS
QASGTEAAIMMQKSNFKGPKRSIKCFNCGKEGHLARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSNKGRPGNFLQNRPEPTAPP
AESFGFGEETAPSPKQEPKEKELYPLASLKSLFGNDP$

Fig. 78B

2003_CON_06_CPX_gag.OPT

ATGGGCGCCCGCGCCTCCGTGCTGTCCGGCGGCAAGCTGGACGAGTGGGAGAAGATCCGCCT
GAAGCACCTGGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCCTGAACCCCGGCCTGCTG
AGCAGCTGCAGTCCGCGACCAGGCCCTGAAGACCGGCTGAAGTCCCTGTACAACACCGTGG
AAGGTGACCGACACCAAGGAGGCCCTGGACAAGATCGAGGAGATCCAGAACAAGTCCAAGC
CAACTCCTCCAACCTGTCCCAGAACTACCCCATCGTGCAGAACGCCCAGGGCCAGATGGTG
CCTGGGTGAAGGTGATCGAGGAGAAGGCCTTCTCCCCGAGGTGATCCCCATGTTCTCCGCC
AACATGATGCTGAACATCGTGGGCGGCCCCATCCCCCCGGCCACCACCCTGTCCACCCTGC
GCACCCCGTGCACGCGCTGATGACCTGCCGCGGCCCCCATCCCCCCCCGGACATCCGCCAG
AGATCGGCTGGATGACCTCCATCCTGGACAGAGGTGAAGAAGCGCTGGATCATCCTGGGCC
TACTCCCCCCGTGTCATCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGT
GCAGGCCACCCAGGCCGCCAACCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATC
GCCCCGGCGCCACCCTGGAGGAGATGATGACCGCGCCTGTGCTGGCCGAGGCCATGTCC
CAGGCCTCCGGCACCGAGGCCGCCATCATGATGCAGAAGTCCAACTTCAAGGGCCCCAAGA
GGTCCATCAAGTGCTTCAACTGCGGCAAGGA
GGGCCACCTGGCCCGCAACTGCCGCGCCCCGCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGA
GGGCCACCAGATGAAGGACTGCACCGAGC
GCCAGGCCAACTTCCTGGGCAAGATCTGGCCTTCCAACAAGGGCCGCCCGGGCAACTTCCTGCAGAACCGCCCCGAGCCCACCGCCCCCCC
CGAGTCCTTCGGCTTCGGCGAGGAGGAGACCGCCCCTCCCCCAAGCAGGAGCCCAAGGAGAAGGAGCTGTACCCCCTGGCCTCCCTGAAGTC
CCTGTTCGGCAACGACCCCTAA

Fig. 79A

20. 2003_CON_07_BC_gag.PEP

MGARASILRGGKLDKWEKIRLRPGGKKHYMLKHLVWASRELERFALNPGLLETSEGCKQIIKQLQPALQTGTEELRSLFNTVATLYCVHTEI
DVRDTKEALDKIEEEQNKIQQKTQQAKEADGKVSQNYPIVQNLQGQMVHQPISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTM
LNTVGGHQAAMQILKDTINEEAAEWDRLHPVHAGPIAPGQMREPRGSDIAGTTSNLQEQIAWMTSNPPVPVGDIYKRWIILGLNKIVRMYSP
TSILDIKQGPKEPFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQNANPDCKTILRALGPGASIEMMTACQGVGGPSHKARVLAEAMSQTN
STILMQRSNFKGSKRIVKCFNCGKEGHIARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSHKGRPGNFLQSRPEPTAPPEESFRF
GEETTTPSQKQEPIDKELYPLTSLKSLFGNDPSSQ$

Fig. 79B

2003_CON_07_BC_gag.OPT

ATGGGCGCCCGCGCCTCCATCCTGCGCGGCGGCAAGCTGGACAAGTGGGAGAAGATCCGCCTGCGCCCCGGCGGCAAGAAGCACTACATGCT
GAAGCACCTGGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCCTGAACCCCGGCCTGCTGGAGACCTCCGAGGGCTGCAAGCAGATCATCA
AGCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCTCCCTGTTCAACACCGTGGCCACCCTGTACTGCGTGCACACCGAGATC
GACGTGCGCGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGAACAAGATCCAGCAGCCCATCTCCCCCCGCACCCTGAACGCCTGGGTGA
CAAGGTGTCCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAGCCCATCTCCCCCCGCACCCTGAACGCCTGGGTGA
AGGTGGTGGAGGAGAAGGCCTTCTCCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCCCAGGACCTGAACACCATG
CTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATCCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCTGCACCCCGT
GCACGCCGGCCCCATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCTCCGACATCGCCGGCACCACCTCCAACCTGCAGGAGCAGATCGCCT
GGATGACCTCCAACCCCCCCGTGCCCGTGGGCGACATCTACAAGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCCCCC
ACCTCCATCCTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCCTGCGCGCCGAGCAGGCCAC
CCAGGACGTGAAGAACTGGATGACCGACACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCCCTGGGCCCCGGCG
CCTCCATCGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCTCGCGTGCTGGCCGAGGCCATGTCCCAGACCAAC
TCCACCATCCTGATGCAGCGCTCCAACTTCAAGGGCTCCAAGCGCATCGTGAAGTGCTTCAACTGCGGCAAGGAGGGCCACCAGATCGCCCGCAA
CTGCCGCGCCCCCAGGAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTGG
GCAAGATCTGGCCCTCCCACAAGGGCCGCCCCGGCAACTTCCTGCAGTCCCGCCCCGAGCCCACCGCCCCCCCCGAGGAGTCCTTCCGCTTC
GGCGAGGAGACCACCACCCCCTCCCAGAAGCAGGAGCCCATCGACAAGGAGCTGTACCCCCTGACCTCCCTGAAGTCCCTGTTCGGCAACGA
CCCCTCCTCCCAGTAA

Fig. 80A

21. 2003_CON_08_BC_gag.PEP

MGARASILRGGKLDKWEKIRLRPGGKKHYMLKHLVWASRELERFALNPGLLETSEGCKQIIKQLQPALQTGTEELRSLFNTVATLYCVHAEI
EVRDTKEALDKIEEEQNKIQQKTQQAKEADEKVSQNYPIVQNLQGQMVHQPLSPRTLNAWVKVVEEKAFSPEVIPMFTALSEGATPQDLNTM
LNTVGGHQAAMQMLKDTINEEAAEWDRLHPVHAGPVAPGQMREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIVRMYSP
TSILDIKQGPKEPFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQNANPDCKTILRALGPGASLEEMMTACQGVGGPSHKARVLAEAMSQTN
NTILMQRSNFKGSKRIVKCFNCGKEGHIAKNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSHKGRPGNFLQSRPEPTAPPAESFRF
EETTPAPKQEPKDREPLTSLRSLFGSDPLSQ$

Fig. 80B

2003_CON_08_BC_gag.OPT

ATGGGCGCCCGCGCCTCCATCCTGCGCGGCGGCAAGCTGGACAAGTGGGAGAAGATCCGCCTGCGCCCTGGGCGGCAAGAAGCACTACATGCT
GAAGCACCTGGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCCTGAACCCCGGCCTCCTGGAGACCTCCGAGGGCTGCAAGCAGATCATCA
AGCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCTCCCTGTTCAACACCGTGGCCACCCTGTACTGCGTGCACGCCGAGATC
GAGGTGCGCGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGAACAAGATCCAGCAGAAGACCCAGCAGGCCAAGGAGGCCGACGA
GAAGGTGTCCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCCCCTGTCCCCCCGCACCCTGAACGCCTGGGTGA
AGGTGGTGGAGGAGAAGGCCTTCTCCCCGAGGTGATCCCCGAGTGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCTGCA
GCACGCCGTGGGCCCCGTGGCCCCCGGCCAGATGCGCGAGCCCCGAGGCTCCGACATCGCTGGAATCATCCTGGCCTGAACAAGATCGTGCGCA
GGATGACCAACCAACCCCCCCCATCCCGTGGGCGAGATCTACAAGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCCCCC
ACCTCCATCCTGGACATCAAGCAGGGCCCAGAACGCTCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCCTGCGCGCCGAGCAGGCCAC
CCAGGACGTGAAGAACTGGATGACCGACCGCTCCTGCTGGGCGGCGCGCCCCCTGGCCCTGGGCCGCG
CCTCCCTGGAGGAGATGATGACCGCCTCCAAGGGCCTCCAAGGCGTGGGCGGCCCTGGCCAAGGAGAAGCCCAGACCAAC
AACACCATCCTGATGCAGCGCTCCAACTTCAAGGGCTCCAAGCGCATCGTGAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAGAA
CTGCCGCGCCCGCCCCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTGG
GCAAGATCTGGCCCTCCCACAAGGGCCGCCCCGGCAACTTCCTGCAGTCCCGCCCCGAGCCCACCGCCCCCCCCGAGTCCTTCCGCTTC
GAGGAGACCACCCCCGCCCCCAAGCAGGAGCCCAAGGACCGCGAGCCCCTGACCTCCCTGCGCTCCCTGTTCGGCTCCGACCCCCTGTCCCCA
GTAA

Fig. 81A

22. 2003_CON_10_CD gag.PEP

MGARASVLSGGKLDEWEKIRLRPGGKKKYRLKHLVWASRELERFALNPGLLETSEGCKQIIGQLQPAIQTGSEEIKSLYNTVATLYCVHERI
KVTDTKEALDKIEEEQTKSKKKAQQATADTGNSSQVSQNYPIVQNLQGQMVHQPLSPRTLNAWVKVIEEKAFSPEVIPMFSALSEGATPQDL
NTMLNTVGGHQAAMQMLKETINEEAAEWDRLHPVQAGPVAPGQIREPRGSDIAGTTSTLQEQIRWMTSNPPIPVGEIYKRWIILGLNKIVRM
YSPVSILDIRQGPKEPFRDYVDRFYKTLRAEQASQDVKNWMTETLLVQNANPDCKTILKALGPAATLEEMMTACQGVGGPSHKARVLAEAMS
QATSGNAIMMQRGNFKGPKKIIKCFNCGKEGHIAKNCRAPRKKGCWKCGREGHQMKDCTERQANFLGKIWPSNKGRPGNFLQSRPEPTAPPA
ESFGFGEEITPSQKQEQKDKELHPLASLKSLFGNDPLSQS

Fig. 81B

2003_CON_10_CD gag.OPT

ATGGGCGCCCGCGCCTCCGTGCTGTCCGGCGGCAAGCTGGACGAGTGGGAGAAGATCCGCCTGCGCCCGGCGGCAAGAAGAAGTACCGCCT
GAAGCACCTGGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCCTGAACCCCGGCCTGCTGGAGACCTCCGAGGGCTGCAAGCAGATCATCG
GCCAGCTGCAGCCCGCCATCCAGACCGGCTCCGAGGAGATCAAGTCCCTGTACAACACCGTGGCCACCCTGTACTGCGTGCACGAGCGCATC
AAGGTGACCGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGACCAAGTCCAAGAAGAAGGCCCAGCAGGCCACCGCCGACACCGG
CAACTCCTCCCAGGTGTCCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAGCCCCTGTCCCCCCGCACCCTGAACG
CCTGGGTGAAGGTGATCGAGGAGAAGGCCTTCTCCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCCAGGACCTG
AACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGAGACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCT
GCACCCCGTGCAGGCCGGCCCCGTGGCCCCCGGCCAGATCCGCGAGCCCCGCGGCTCCGACATCGCTGGCACTACCTGGCCTTCTACAAGC
AGATCCGCTGGATGACCTCCAATCCTGGACATCCGCGAGGGCGCCCAAGGAGCCCTTCCGCAGACCCCGACTGTGGTGCAGAACGCCAACCCCGA
TACTCCCCCGTGTCCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTACAAGACCCTGCGCGCCGA
GCAGGCCTCCCAGGACGTGAAGAACTGGATGACCGAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCCCTGG
GCCCCGCCGCCACCCTGGAAGAAATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTGCTGGCCGAGGCCATGTCC
CAGGCCACCTCCGGCAACGCCATCATGATGCAGCGCGGCAACTTCAAGGGCCCCAAGAAGATCATCAAGTGCTTCAACTGCGGCAAGGAGGG
CCACATCGCCAAGAACTGCCGCGCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCCGCGAGGGCCACCAGATGAAGGACTGCACCGAGCGCC
AGGCCAACTTCCTGGGCAAGATCTGGCCCTCCAACAAGGGCCGCCCCGGCAACTTCCTGCAGTCCCGCCCCGAGCCCACCGCCCCCCCCGCC
GAGTCCTTCGGCTTCGGCGAGGAGATCACCCCCTCCCAGAAGCAGGAGCAGAAGGACAAGGAGCTGCACCCCCTGGCCTCCCTGAAGTCCCT
GTTCGGCAACGACCCCCTGTCCCAGTAA

Fig. 82A

23. 2003_CON_11_CPX gag.PEP

gag.PEPMGARASVLSGGKLDAWEKIRLRPGGKKKYRLKHLVWASRELERFALNPSLLETAEGCQQIMGQLQPALGTGTEELRSLYNTVATL
YCVHHRIEVKDTKEALDKIEEIQNKSKQKKQQAAADTGNSSKVSQNYPIVQNAGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSE
GATPQDLNMMLNIVGGHQAAMQMLKDTINEEAAEWDRVHPVHAGPIPPGQMREPRGSDIAGTTSTLQEQIGWMTGNPPVPVGEIYRRWIILG
LNKIVRMYSPVSILDIRQGPKEPFRDYVDRFFKTLRAEQATQEVKSWMTETLLIQNANPDCKSILRALGPGATLEEMMTACQGVGGPGHKAR
VLAEAMSQVQQTNIMMQRSNFKGQKRIKCFNCGKEGHLARNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSSKGRPGNFLQSRPEP
TAPPAESFGFGEEIAPSPKQEPKEKELYPLTSLKSLFGSDPLSQ$

Fig. 82B

2003_CON_11_CPX gag.OPT

ATGGGCGCCCGCGCCTCCGTGCTGTCCGGCGGCAAGCTGGACGCCTGGGAGAAGATCCGCCTGCGCCTGCGCCTGAAGCACCTGGTGTGGG
GAAGCACCTGGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCCTGAACCCCTCCCTGCTGGAGACCGCCGAGGGCTGCCAGCAGATCATGG
GCCAGCTGCAGCCCGCCCTGGGCACCGGCACCGAGGAGCTGCGCTCCCTGTACAACACCGTGGCCACCCTGTACTGCGTGCACCACCGCATC
GAGGTGAAGGACACCAAGGAGGCCCTGGACAAGATCGAGGAGATCCAGAACAAGTCCAAGCAGAAGAAGCAGCAGGCCGCCGCCGACACCGG
CAACTCCTCCAAGGTGTCCCAGAACTACCCCATCGTGCAGAACGGCCAGGGCCATGTGCACCAGGCCATCTCCCCGCCACCCTGAACG
CCTGGGTGAAGGTGGTGGAGGAGAAGGCCTTCTCCCCGGAGGTGATCCAGATGTTCTCCGCCCTGAGTGAACGAGGGCCACCCCCCAGGACCTG
AACATGATGCTGAACATCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGT
GCACCCCGTGCACGCCGGCCCCATCCCCCCCGGCCAGATGCGCGAACCGCGGAGCCCCGCGAGGCGTGGGCGTGGGCGGCGAGATCTACCGCCGCTGGATC
AGATCGGCTGGATGACCGGCAACCCCCCGGGTCCCGTGCCCGTGGGCGAGATCTACCGCCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATG
TACTCCCCCGTGTCCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCCTGCGCGCGGA
GCAGGCCGAGCAGGCCACCCAGGAGGTGAAGTCCTGGATGACCGAGACGCTGCTGATCCAGAACGCCAACCCCGACTGCAAGTCCATCCTGCGCGCCCTGG
GCCCCGGCGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTGCTGGCCGAGGCCATGTCC
CAGGTGCAGCAGACCAACATCATGATGCAGCGCTCCAACTTCAAGGGCCAGAAGCGCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACCT
GGCCCGCAACTGCCGCGCCCCCGCAGAAGGGGCTGCTGGAAGTGCGGCCCCGGCAACTTCCTGCAGTCCCGCCCCGAGCCCCCCGCGAGTCC
ACTTCCTGGCCAAGATCTGGCCCTCCTCCAAGGGCCCCCGCCCCAAGGAGCAGGAGAAGGAGCTGTACCCCCTGACCTCCCTGAAGTCCCTGTTCGG
CTCCGACCCCCTGTCCCAGTAA

Fig. 83A

24. 2003_CON_12_BF.gag.PEP

MGARASVLSGGELDRWEKIRLRPGGKKKYRLKHIVWASRELERFAVNPGLLETSEGCRKIIGQLQPSLQTGSEELRSLYNTIAVLYFVHQKV
EVKDTKEALDKLEEEQNKSQQKTQQAAADKGVSQNYPIVQNLQGQMVHQALSPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTML
NTVGGHQAAMQMLKDTINEEAAEWDRLHPVHAGPIPPGQMREPRGSDIAGTTSTLQEQIQWMTSNPPVPVGEIYKRWIILGLNKIVRMYSPV
SILDIRQGPKEPFRDYVDRFFKTLRAEQATQEVKGWMTDTLLVQNANPDCKTILKALGPGATLEEMMTACQGVGGPGHKARVLAEAMSQVTN
TTVMMQKSNFKGQRRIVKCFNCGKEGHIAKNCRAPRKKGCWKCGREGHQMKDCTERQANFLGKIWPSNKGRPGNFLQNRPEPTAPPAESFGF
GEEITPSPKQEQKDEGLYPPLASLKSLFGNDP$

Fig. 83B

2003_CON_12_BF.gag.OPT

ATGGGCGCCCGCGCCTCCGTGCTGTCCGGCGGCGAGCTGGACCGCTGGGAGAAGATCCGCCTGCGCCCCGGCGGCAAGAAGAAGTACCGCCT
GAAGCACATCGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCGTGAACCCCGGCCTCCTGGAGACCTCCGAGGGCTGCCGCAAGATCATCG
GCCAGCTGCAGCCCTCCCTGCAGACCGGCCTGGAGGAGCTGCGCTCCCTGTACAACACCATCGCCGTGCTGTACTTCGTGCACCAGAAGGTG
GAGGTGAAGGACACCAAGGAGGCCCTGGACAAGCTGGAGGAGGAGCAGAACAAGTCCCAGCAGAAGACCCAGCAGGCCGCCGCCGACAAGGG
CGTGTCCCAGAACTACCCCATCGTGCAGAACCTGCAGGGTGCAGGGCCAGATGGTGCACCAGGCCCTGTCCCCCCGCACCCTGAACGCCTGG
TGGTGAAGGTGGTGGAGGAGAAGGCCTTCTCCCCCGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCCCAGGACCTGCTG
AACACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCTG
CGCCCCGTGCACGCCGGCCCCATCCCCCCCGGCCAGATGCGCGAGCCCCGCGGCTCCGACATCGCCGGCACCACCTCCACCCTGCAGGAGCAGATCCAGTGGA
TGACCTCCAACCCCCCCGTGCCCGTGGGCGAGATCTACAAGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACTCCCCCGTG
TCCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCATCCTGCGCGCCGAGCAGGCCACCCA
GGAGGTGAAGGGCTGGATGACCGACACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGAAGGCCCTGGGCCCCGGCGCCA
CCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCAGGGCCATGTCCCAGGTGACCAAC
ACCACCGTGATGATGCAGAAGTCCAACTTCAAGGGCCAGCGCCGCATCGTGAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAGAA
CTGCCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCCGCGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCCAGGCCCAACTTCCTGG
GCAAGATCTGGCCCTCCAACAAGGGCCGCCCCGGCAACTTCCTGCAGAACCGCCCCGAGCCCACCGCCCCCGCCGAGTCCTTCGGCTTC
GGCGAGGAGATCACCCCCTCCCCCAAGCAGGAGCAGAAGGACGAGGGCCTGTACCCCCCCCTCCCTGAAGTCCCTGTTCGGCAACGA
CCCCTAA

SUBSTITUTE SHEET (RULE 26)

Fig. 84A

25. 2003 CON 14 BG gag.PEP

MGARASVLSGGKLDAWEKIRLRPGGKKKYRMKHLVWASRELERFALNPDLLETAEGCQQIMGQLQPALQTGTEEIRSLFNTVATLYCVHQKI
EVKDTKEALEEVEKAQKKSQKKQQAAMDEGNNSQASQNYPIVQNAQGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLN
TMLNTVGGHQAAMQMLKDTINEEAAEWDRMHPQQAGPIPPGQIREPRGSDIAGTTSTLQEQIRWMTSNPPIPVGEIYKRWIILGLNKIVRMY
SPVSILDIRQGPKEPFRDYVDRFFKTLRAEQATQEVKGWMTDTLLVQNANPDCKTILRALGPGATLEEMTACQGVGGPSHKARVLAEAMSQ
ASGATIMMQKSNFKGPRRNIKCFNCGKEGHLARNCRAPRKKGCWKCGKEGHQMKDCTESKANFLGKIWPSNKGRPGNFLQNRPEPTAPPAES
FGFGEEIAPSPKQEPKEKEIYPLASLKSLFGSDPSQ

Fig. 84B

2003 CON 14 BG gag.OPT

ATGGGCGCGCGCCTCCGTGCTGTCCGGCGGCAAGCTGGACGCGCCTGGGAGAAGATCCGCCTGCGCCTGCCCGGCGGCAAGAAGAAGTACCGCAT
GAAGCACCTGGTGTGGGCCTCCCGCGAGCTGGAGCGCTTCGCCCTGAACCCCGACCTGCTGGAGACCGCCGAGGGCTGCCAGCAGATCATGG
GCCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGATCCGCTCCCTGTTCAACACCGTGGCCACCCTGTACTGCGTGCACCAGAAGATC
GAGGTGAAGGACACCAAGGAGGCCCTGGAGGAGGTGGAGAAGGCCCAGAAGAAGTCCCAGAAGAAGCAGCAGGCCGCCATGGACGAGGGCAA
CAACTCCCAGGCCTCCCAGAACTACCCCATCGTGCAGAACGCCCAGGGCCAGATGGTGCACCAGGCCATCTCCCCGCGCACCCTGAACGCCT
GGGTGAAGGTGGTGGAGGAGAAGGCCTTCTCCCCGGAGGTGATCCCCATGTTCTCCGCCCTGTCCGAGGGCGCCACCCCCAGGACCTGAAC
ACCATGCTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCATGCA
CCCCCAGCAGGCCGGCCCCATCCCCCCCGGCCAGATCCGCGAGCCCCGTGGGCCGGCGACATCGCCGGCACCACCTCCACCCTGCAGGAGCAGA
TCCGCTGGATGACCTCCAACCCCCCCATCCCCGTGGGCGAGATCTACAAGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTAC
TCCCCCGTGTCCATCCTGGACATCCGCCAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCCTGCGCGCCGAGCA
GGCCACCCAGGAGGTGAAGGGCTGGATGACCGACCTGCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGCGCGCCCTGGGCC
CCGGCGCCACCCTGGAGGAGATGACCGCCTGCCAGGGCGTGGGCGGCCCCTCCCACAAGGCCCGCGTGCTGGCCGAGGCCATGTCCCAG
GCCTCCGGCGCCACCATCATGATGCAGAAGTCCAACTTCAAGGGCCCGCGCCGCAACATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACCT
GGCCCGCAACTGCCGCGCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGTCCAAGGCCA
ACTTCCTGGGCAAGATCTGGCCCTCCAACAAGGGCCGCCCCGGCAACTTCCTGCAGAACCGCCCCGAGCCCACCGCCCCCCCCGAGTCC
TTCGGCTTCGGCGAGGAGATCGCCCCCTCCCCCAAGCAGGAGCCCAAGGAGAAGGAGATCTACCCCCTGGCCTCCCTGAAGTCCCTGTTCGG
CTCCGACCCCTAATCCCCAGTAA

Fig. 85A

31. 2003_CONS nef.PEP

MGGKWSKSSIVGWPAVRERIRRTPPAAEGVGAVSQDLDKHGAITSSNTAATNADCAWLEAQEEEVGFPVRPQVPLRPMTYKGAFDLSHFLK
EKGGLDGLIYSKKRQEILDLWVYHTQGYFPDWQNYTPGPGIRYPLTFGWCFKLVPVDPEEVEEANEGENNCLLHPMCQHGMEDEDREVLMWK
FDSRLALRHIARELHPEFYKDC$

Fig. 85B

2003_CONS nef.OPT

ATGGGCGGCAAGTGGTCCAAGTCCTCCATCGTCGGCTGGCCCGCCGTGCGCGAGCGCATCCGCCGCACCCCCCCCGCCGAGGGCGTGGG
CGCCGTGTCCCAGGACCTGGACAAGCACGGCGCCATCACCTCCTCCAACACGGCCGCCACCAACGCCGACTGCGCCTGGCTGGAGGCCCAGG
AGGAGGAGGTGGGCTTCCCCGTGCGCCCCCAGGTGCCCCTGCGCCCCATGACCTACAAGGGCGCCTTCGACCTGTCCCACTTCCTGAAG
GAGAAGGGCGGCCTGGACGGCCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACCACACCCAGGGCTACTTCCCCGA
CTGGCAGAACTACACCCCCGGCCCCGGCATCCGCTACCCCCTGACCTTCGGCTGGTGCTTCAAGCTGGTGCCCGTGGACCCCGAGGAGGTGG
AGGAGGCCAACGAGGGCGAGAACAACTGCCTGCTGCACCCCATGTGCCAGCACGGCATGGAGGACGAGGACCGCGAGGTGCTGATGTGGAAG
TTCGACTCCCGCCTGGCCCTGCGCCACATCGCCCGCGAGCTGCACCCCGAGTTCTACAAGGACTGCTAA

Fig. 86A

32. 2003_M. GROUP.anc nef.PEP

MGGKWSKSSIVGWPAVRERMRRTAPAAEGVGAVSQDLDKHGAITSSNTAATNADCAWLEAQEEEVGFPVRPQVPLRPMTYKAAFDLSHFLK
EKGGLDGLIYSKKRQEILDLWVYHTQGYFPDWQNYTPGPGIRYPLTFGWCFKLVPVDPEEVEEANEGENNCLLHPMCQHGMEDEEREVLMWK
FDSRLALRHIARELHPEFYKDC$

Fig. 86B

2003_M GROUP.anc nef.OPT

ATGGGCGGCAAGTGGTCCAAGTCCTCCATCGTCGGCTGGCCCGCCGTGCGCGAGCGCATGCGCCGCACCGCCCCCGCCGCCGAGGGCGTGGG
CGCCGTGTCCCAGGACCTGGACAAGCACGGCGCCATCACCTCCTCCAACACGGCCGCCACCAACGCCGACTGCGCCTGGCTGGAGGCCCAGG
AGGAGGAGGTGGGCTTCCCCGTGCGCCCCCAGGTGCCCCTGCGCCCCATGACCTACAAGGCCGCCTTCGACCTGTCCCACTTCCTGAAG
GAGAAGGGCGGCCTGGACGGCCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACCACACCCAGGGCTACTTCCCCGA
CTGGCAGAACTACACCCCCGGCCCCGGCATCCGCTACCCCCTGACCTTCGGCTGGTGCTTCAAGCTGGTGCCCGTGGACCCCGAGGAGGTGG
AGGAGGCCAACGAGGGCGAGAACAACTGCCTGCTGCACCCCATGTGCCAGCACGGCATGGAGGACGAGGAGCGCGAGGTGCTGATGTGGAAG
TTCGACTCCCGCCTGGCCCTGCGCCACATCGCCCGCGAGCTGCACCCCGAGTTCTACAAGGACTGCTAA

Fig. 87A

33. 2003 CON A nef.PEP

MGGKWSKSSIVGWPDIRERIRRTPPAAKGVGAVSQDLDKYGAVTINNTAATQASCAWLEAQEEEEVGFPVRPQVPLRPMTFKGAFDLSFFL
KEKGGLDGLIYSQKRQEILDLWVYNTQGYFPDWQNYTPGPGTRFPLTFGWCFKLVPVDPDEVEEATEGENNCLLHPICQHGMDDEEKEVLMW
KFDSRLARRHIALEMHPEFYKDC$

Fig. 87B

2003 CON A nef.OPT

ATGGGCGGCAAGTGGTCCAAGTCCTCCATCGTGGGCTGGCCCGACATCCGCGAGCGCATCCGCCGCACCCCCCCGGCCGCCAAGGGCGTGGG
AGGAGGAGGAGGAGGTGGCTTCCCCGTGCGCCGTGCCCCGAGTGCCCGTGCCCCATGACCTTCAAGGGCGCCTTCGACCTGTCCTTCTTCCTG
AAGGAGAAGGGCGGCCTGGACGGCCTGATCTACTCCCAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACAACACCCAGGGCTACTTCCC
CGACTGGCAGAACTACACCCCCGGCCCCGGCACCCGCTTCCCCCTGACCTTCGGCTGGTGCTTCAAGCTGGTGCCCGTGGACCCCGACGAGG
TGGAGGAGGCCACCGAGGGCGAGAACAACTGCCTGCTGCACCCCATCTGCCAGCACGGCATGGACGACGAGGAAGGAGGTGCTGATGTGG
AAGTTCGACTCCCGCCTGGCCCGCCGCCACATCGCCCTGGAGATGCACCCCGAGTTCTACAAGGACTGCTAA

Fig. 88A

34. 2003 CON A1 nef.PEP

MGGKWSKSSIVGWPEVRERMRRTPPAATGVGAVSQDLDKHGAVTSSNINHPSCVWLEAQEEEEVGFPVRPQVPLRPMTYKGALDLSHFLKEK
GGLDGLIYSRKRQEILDLWVYHTQGYFPDWQNYTPGPGIRYPLTFGWCFKLVPVDPDEVEKATEGENNSLLHPICQHGMDDEEREVLKWKFD
SRLALKHRAQELHPEFYKDC$

Fig. 88B

2003 CON A1 nef.OPT

ATGGGCGGCAAGTGGTCCAAGTCCTCCATCGTGGGCTGGCCCGAGGTGCGCGAGCGCATGCGCCGCACCCCCCCCGCCGCCACCGGCGTGGG
AGGAGGAGGAGGTGGCGCCGTGCCCCCGAGGTGCGCGAGCGCATGCGCCGCACCCCTCCGCGTGTGGCTGGAGGCCCAGGAGGAGG
AGGAGGAGGTGGGCTTCCCCGTGCGCCCTGCCCCGAGTGCGCCAGGTGCCCCTGCGCCCCATGACCTACAAGGGCGCCCTGGACCTGTCCCACTTCCTGAAGGAGAAG
GGCGGCCTGGACGGCCTGATCTACTCCCGCAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACCACACCCAGGGCTACTTCCCCGACTGGCA
GAACTACACCCCCGGCCCCGGCATCCGCTACCCCCTGCTGCACCCCATCTGCCAGCACGGCATGGACGACGAGGAGCGCGAGGTGCTGAAGAGG
TCCCGCCTGGCCCTGAAGCACCGCGCCCAGGAGCTGCACCCCGAGTTCTACAAGGACTGCTAA

Fig. 88C

35. 2003_A1.anc nef.PEP

MGGKWSKSSIVGWPEVRERMRRTPPAAKGVGAVSQDLDKHGAVTSSNTAANNPGCAWLEAQEEEVGFPVRPQVPLRPMTYKGAFDLSHFLK
EKGGLDGLIYSKKRQEILDLWVYHTQGYFPDWQNYTPGPGIRYPLTFGWCFKLVPVDPAEVEEATEGENNSLLHPICQHGMDDEEREVLMWK
FDSRLALKHRARELHPEFYKDC$

Fig. 88D

2003_A1.anc nef.OPT

ATGGGCGGCAAGTGGTCCAAGTCCTCCATCGTGGGCTGGCCCGAGGTGCGCGAGCGCATGCGCCGCACCCCCCGCCGCCAAGGGCGTGGG
CGCCGGTGTCCCAGGACCTGGACAAGCACGGCGCCGTGACCTCCTCCAACACGCCGCCAACAACCCGCCTGCGCCTGCGGCTGCGAGGCCCAGG
AGGAGGAGGAGGTGGGCTTCCCCGTGCGCCCCCAGGTGCCCCTGCGCCCCATGACCTACAAGGGCGCCTTCGACCTGTCCCACTTCCTGAAG
GAGAAGGGCGGCCTGGACGGCCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACCACACCCAGGGCTACTTCCCCGA
CTGGCAGAACTACACCCCCGGCCCCGGCATCCGCTACCCCCTGACCTTCGGCTGGTGCTTCAAGCTGGTGCCCGTGGACCCCGCCGAGGTGG
AGGAGGCCACCGAGGGCGAGAACAACTCCCTGCTGCACCCCATCTGCCAGCACGGCATGGACGACGAGGAGCGCGAGGTGCTGATGTGGAAG
TTCGACTCCCGCCTGGCCCTGAAGCACCGCGCCCGCGAGCTGCACCCCGAGTTCTACAAGGACTGCTAA

Fig. 89A

36. 2003_CON_A2 nef.PEP

MGGKWSKSSIVGWPAIRERMRKRTPPAAEGVGAVSQDLATRGAVTSSNTAATNPDCAWLEAQEEEVGFPVRPQVPLRPMTFKGAFDLSHFL
KEKGGLDGLIYSQKRQDILDLWVYHTQGYFPDWQNYTPGPGTRYPLTFGWCFKLVPVDPSEVEEATEGENNSLLHPICQHGIEDPEREVLRW
KFDSRLALRHRARELHPEFYKDC$

Fig. 89B

2003_CON_A2 nef.OPT

ATGGGCGGCAAGTGGTCCAAGTCCTCCATCGTGGGCTGGCCCGCCATCCGCGAGCGCATGCGCAAGCGCACCCCCCCGCCGCCGAGGGCGT
GGGCGCCGGTGTCCCAGGACCTGGCCACCCGCGGCGCCGTGACCTCCTCCAACACCGCCGCCACCAACCCCGACTGCGCCTGGCTGGAGGCCC
AGGAGGAGGAGGTGGGCTTCCCCGTGCGCCCCCAGGTGCCCCTGCGCCCCATGACCTTCAAGGGCGCCTTCGACCTGTCCCACTTCCTG
AAGGAGAAGGGCGGCCTGGACGGCCTGATCTACTCCCAGAAGCGCCAGGACATCCTGGACCTGTGGGTGTACCACACCCAGGGCTACTTCCC
CGACTGGCAGAACTACACCCCCGGCCCCGGCACCCGCTACCCCCTGACCTTCGGCTGGTGCTTCAAGCTGGTGCCCGTGGACCCCTCCGAGG
TGGAGGAGGCCACCGAGGGCGAGAACAACTCCCTGCTGCACCCCATCTGCCAGCACGGCATCGAGGACCCCGAGCGCGAGGTGCTGCGCTGG
AAGTTCGACTCCCGCCTGGCCCTGCGCCACCGCGCCCGCGAGCTGCACCCCGAGTTCTACAAGGACTGCTAA

SUBSTITUTE SHEET (RULE 26)

Fig. 90A

37. 2003_CON_B nef.PEP

MGGKWSKRSVVGWPTVRERMRRAEPAADGVGAVSRDLEKHGAITSSNTAANNADCAWLEAQEEEVGFPVRPQVPLRPMTYKGALDLSHFLK
EKGGLEGLIYSQKRQDILDLWVYHTQGYFPDWQNYTPGPGIRYPLTFGWCFKLVPVEPEKVEEANEGENNSLLHPMSLLHGMDDPEREVLVWK
FDSRLAFHHMARELHPEYYKDC$

Fig. 90B

2003 CON-B nef.OPT

ATGGGCGGCAAGTGGTCCAAGCGCTCCGTGGTGGGCTGGCCCACCGTGCGCGAGCGCATGCGCCGCGCCGAGCCCGCAGCGGCGTGGG
CGCCGTGTCCCGCGACCTGGAGAAGCACGGCGCCATCACCTCCTCCAACACCGCCGCCAACAACGCCGACTGCGCCTGGCTGGAGGCCCAGG
AGGAGGAGGAGGTGGGCTTCCCCGTGCGCCCCCAGGTGCCCCTGCGCCCCATGACCTACAAGGGCGCCCTGGACCTGTCCCACTTCCTGAAG
GAGAAGGGCGGCCTGGAGGGCCTGATCTACTCCCAGAAGCGCCAGGACATCCTGGACCTGTGGGTGTACCACACCCAGGGCTACTTCCCCGA
CTGGCAGAACTACACCCCCGGCCCCGGCATCCGCTACCCCCTGACCTTCGGCTGGTGCTTCAAGCTGGTGCCCGTGGAGCCCGAGAAGGTGG
AGGAGGCCAACGAGGGCGAGAACAACTCCCTGCTGCACCCCATGTCCCTGCTGCACGGCATGGACGACCCCGAGAGCGAGGTGCTGGTGTGGAAG
TTCGACTCCCGCCTGGCCTTCCACCACATGGCCCGCGAGCTGCACCCCGAGTACTACAAGGACTGCTAA

Fig. 90C

38. 2003_B.anc nef.PEP

MGGKWSKSSMGGWPAVRERMKRAEPAADGVGAVSRDLEKHGAITSSNTAATNADCAWLEAQEEEVGFPVRPQVPLRPMTYKAALDLSHFLK
EKGGLEGLIYSQKRQDILDLWVYHTQGYFPDWQNYTPGPGIRYPLTFGWCFKLVPVEPEKVEEATEGENNSLLHPMSLLHGMDDPEKEVLVWK
FDSRLAFHHMARELHPEYYKDC$

Fig. 90D

2003 B.anc nef.OPT

ATGGGCGGCAAGTGGTCCAAGTCCTCCATGGGCGGCTGGCCCGCCGTGCGCGAGCGCATGAAGCGCGCCGAGCCCGCCGCCGACGGCGTGGG
CGCCGTGTCCCGCGACCTGGAGAAGCACGGCGCCATCACCTCCTCCAACACCGCCGCCACCAACGCCGACTGCGCCTGGCTGGAGGCCCAGG
AGGAGGAGGTGGGCTTCCCCGTGCGCCCCCAGGTGCCCCTGCGCCCCATGACCTACAAGGCCGCCCTGGACCTGTCCCACTTCCTGAAG
GAGAAGGGCGGCCTGGAGGGCCTGATCTACTCCCAGAAGCGCCAGGACATCCTGGACCTGTGGGTGTACCACACCCAGGGCTACTTCCCCGA
CTGGCAGAACTACACCCCCGGCCCCGGCATCCGCTACCCCCTGCTGCTGCACCCCATGTGCCTGCTGCACGGCATGGACGACCCCGAGAAGGTGG
AGGAGGCCACCGAGGGCGAGAACAACTCCCTGCTGCACCCCATGTCCCTGCTGCACGGCATGGACGACCCCGAGAAGGAGGTGCTGGTGTGGAAG
TTCGACTCCCGCCTGGCCTTCCACCACATGGCCCGCGAGCTGCACCCCGAGTACTACAAGGACTGCTAA

Fig. 91A

39. 2003_CON_02_AG nef.PEP

MGGKWSKSSIVGWPKVRERIRQTPPAATGVGAASQDLDRHGAITSSNTAATNADCAWLEAQEEEVGFPVRPQVPLRPMTYKAAVDLSHFLK
EKGGLEGLIYSKKRQEILDLWVYHTQGFFPDWQNYTPGPGTRFPLTFGWCFKLVPMDPAEVEEANEGENNSLLHPICQHGMEDEDREVLVWR
FDSSLAFKHRARELHPEFYKDC$

Fig. 91B

2003_CON_02_AG nef.OPT

ATGGGCGGCAAGTGGTCCAAGTCCTCCATCGTGGGCTGGCCCAAGGTGCGCGAGCGCATCCGCCAGACCCCCCCCGCCGCCACCGGCGTGGG
CGCCGCCTCCCAGGACCTGGACCGCCATCACCTCCTCCAACACCGCCGCCACCAACGCCGACTGCGCCTGGCTGGAGGCCCAGG
AGGAGGAGGAGGTGGGCTTCCCCGTGCGCCCCCAGGTGCCCCTGCGCCCCATGACCTACAAGGCGCGGTGACCACACCCAGGGCTTCTTCCCCGA
GAGAAGGGCGGCCTGGAGGGCCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACCACACCCAGGGCTTCTTCCCCGA
CTGGCAGAACTACACCCCCGGCCCCGGCACCCGCTTCCCCCTGACCTTCGGCTGGTGCTTCAAGCTGGTGCCCATGGACCCCGCCGAGGTGG
AGGAGGCCAACGAGGGCGAGAACAACTCCCTGCTGCACCCCATCTGCCAGCACGGCATGGAGGACGAGGACCGCGAGGTGCTGGTGTGGCGC
TTCGACTCCTCCCTGGCCTTCAAGCACCGCGCCCGCGAGCTGCACCCCGAGTTCTACAAGGACTGCTAA

Fig. 92A

40. 2003_CON_C nef.PEP

MGGKWSKSSIVGWPAVRERIRRTEPAAEGVGAASQDLDKHGALTSSNTATNNADCAWLEAQEEEEVGFPVRPQVPLRPMTYKAAFDLSFFL
KEKGGLEGLIYSKKRQEILDLWVYHTQGYFPDWQNYTPGPGVRYPLTFGWCFKLVPVDPREVEEANEGENNCLLHPMSQHGMEDEDREVLKW
KFDSHLARRHMARELHPEYYKDC$

Fig. 92B

2003_CON_C nef.OPT

ATGGGCGGCAAGTGGTCCAAGTCCTCCATCGTGGGCTGGCCCGCCGTGCGCGAGCGCATCCGCCGCACCGAGCCCGCCGCCGAGGGCGTGGG
CGCCGCCTCCCAGGACCTGGACAAGCACGGCGCCCTGACCTCCTCCAACACCGCCACCAACAACGCCGACTGCGCCTGGCTGGAGGCCCAGG
AGGAGGAGGAGGTGGGCTTCCCCGTGCGCCCCCAGGTGCCCCTGCGCCCCATGACCTACAAGGCCGCCTTCGACCTGTCCTTCTTCCTG
AAGGAGAAGGGCGGCCTGGAGGGCCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACCACACCCAGGGCTACTTCCC
CGACTGGCAGAACTACACCCCCGGCCCCGGCGTGCGCTACCCCCTGACCTTCGGCTGGTGCTTCAAGCTGGTGCCCGTGGACCCCCGCGAGG
TGGAGGAGGCCAACGAGGGCGAGAACAACTGCCTGCTGCACCCCATGTCCCAGCACGGCATGGAGGACGAGGACCGCGAGGTGCTGAAGTGG
AAGTTCGACTCCCACCTGGCCCGCCGCCACATGGCCCGCGAGCTGCACCCCGAGTACTACAAGGACTGCTAA

Fig. 92C

41. 2003 C.anc nef.PEP

MGGKWSKSSIVGWPAVRERMRRTEPAAEGVGAASQDLDKHGALTSSNTAANNADCAWLEAQEEEEVGFPVRPQVPLRPMTYKAAFDLSFFL
KEKGGLDGLIYSKKRQEILDLWVYHTQGYFPDWQNYTPGPGVRYPLTFGWCFKLVPVDPREVEEANEGENNCLLHPMSQHGMEDEDREVLKW
KFDSHLARRHMARELHPEYYKDC$

Fig. 92D

2003 C.anc nef.OPT

ATGGGCGGCAAGTGGTCCAAGTCCTCCATCGTGGGCTGGCCCGCCGTGCGCGAGCGCATGCGCCGCACCGAGCCCGCCGCAGAGGGCGTGGG
CGCCGCCTCCCAGGACCTGGACAAGCACGGCGCCCTGACCTCCTCCAACACCGCCGCCAACAACGCCGACTGCGCCTGGCTGGAGGCCCAGG
AGGAGGAGGAGGAGGTGGGCTTCCCCGTGCGCCCCCAGGTGCCCCTGCGCCCCATGACCTACAAGGCCGCCTTCGACCTGTCCTTCTTCCTG
AAGGAGAAGGGCGGCCTGGACGGCCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACCACACCCAGGGCTACTTCCC
CGACTGGCAGAACTACACACCCGGCCCCGGCGTGCGCTACCCCCTGACCTTCGGCTGGTGCTTCAAGCTGGTGCCCGTGGACCCCCGCGAGG
TGGAGGAGGCCAACGAGGGCGAGAACAACTGCCTGCTGCACCCCATGTCCCAGCACGGCATGGAGGACGAGGACCGCGAGGTGCTGAAGTGG
AAGTTCGACTCCCACCTGGCCCGCCGCCACATGGCCCGCGAGCTGCACCCCGAGTACTACAAGGACTGCTAA

Fig. 93A

42. 2003 CON D nef.PEP

MGGKWSKSSIVGWPAIRERIRRTEPAADGVGAVSRDLEKHGAITSSNTAATNADCAWLEAQEEDEEVGFPVRPQVPLRPMTYKAALDLSHFL
KEKGGLEGLVWSQKRQEILDLWVYNTQGFFPDWQNYTPGPGIRYPLTFGWCFELVPVDPEEVEEATEGENNCLLHPMCQHGMEDPEREVLMW
RFNSRLAFEHKARVLHPEFYKDC$

Fig. 93B

2003 CON D nef.OPT

ATGGGCGGCAAGTGGTCCAAGTCCTCCATCGTGGGCTGGCCCGCCATCCGCGAGCGCATCCGCCGCACCGAGCCCGCCGCAGACGGCGTGGG
CGCCGTGTCCCGCGACCTGGAGAAGCACGGCGCCATCACCTCCTCCAACACCGCCGCCACCAACGCCGACTGCGCCTGGCTGGAGGCCCAGG
AGGAGGACGAGGAGGTGGGCTTCCCCGTGCGCCCCCAGGTGCCCCTGCGCCCCATGACCTACAAGGCCGCCCTGGACCTGTCCCACTTCCTG
AAGGAGAAGGGCGGCCTGGAGGGCCTGGTGTGGTCCCAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACAACACCCAGGGCTTCTTCCC
CGACTGGCAGAACTACACACCCGGCCCCGGCATCCGCTACCCCCTGACCTTCGGCTGGTGCTTCGAGCTGGTGCCCGTGGACCCCGAGGAGG
TGGAGGAGGCCACCGAGGGCGAGAACAACTGCCTGCTGCACCCCATGTGCCAGCACGGCATGGAGGACCCCGAGCGCGAGGTGCTGATGTGG
CGCTTCAACTCCCGCCTGGCCTTCGAGCACAAGGCCCGCGTGCTGCACCCCGAGTTCTACAAGGACTGCTAA

Fig. 94A

43. 2003_CON_F1 nef.PEP

MGGKWSKSSIVGWPAVRERMRPTPPAAEGVGAVSQDLERRGAITSSNTGATNPDLAWLEAQEEEVGFPVRPQVPLRPMTYKGAVDLSHFLK
EKGGLEGLIYSKKRQEILDLWVYHTQGYFPDWQNYTPGPGIRYPLTFGWCFKLVPVDPEEVEKANEGENNCLLHPMSQHGMEDEDREVLIWK
FDSRLALRHIARERHPEFYQD$

Fig. 94B

2003_CON_F1 nef.OPT

ATGGGCGGCAAGTGGTCCAAGTCCTCCATCGTGGGCTGGCCCGCCGTGCGCGAGCGCATGCGCCCCACCCCCCCGGCCGCCGAGGGCGTGGG
CGCCGTGTCCCAGGACCTGGAGCGCCGCGGCGCCATCACCTCCTCCAACACCGGCGCCACCAACCCCGACCTGGCTGGCTGGAGCCCAGG
AGGAGGAGGAGGTGGGCTTCCCCGTGCGCCCCCAGGTGCCCCTGCGCCCCATGACCTACAAGGGCGCCGTGGACCTGTCCCACTTCCTGAAG
GAGAAGGGCGGCCTGGAGGGCCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTTCGGCTGGTGTACCACACCCAGGGCTACTTCCCCGA
CTGGCAGAACTACACACCCCCGGCGCATCCCGGCGGCCCCCCCTGACCTTCGGCTGGTGCTTCAAGCTGGTGCCCGTGGACCCCGAGGAGGTGG
AGAAGGCCAACGAGGGCGAGAACAACTGCCTGCTGCACCCCATGTCCCAGCACGGCATGGAGGACGAGGACCGCGAGGTGCTGATCTGGAAG
TTCGACTCCCGCCTGGCCCTGCGCCACATCGCCCGCGAGCGCCACCCCGAGTTCTACCAGGACTAA

Fig. 95A

44. 2003_CON_F2 nef.PEP

MGGKWSKSSIVGWPTIRERIRRTPVAAEGVGAVSQDLDKHGAITSSNTRATNADLAWLEAQEDEEVGFPVRPQVPLRPMTYKAAFDLSHFLK
EKGGLEGLIYSKKRQEILDLWVYHTQGYFPDWQNYTPGPGTRYPLTFGWCFKLVPVDPEEVEKANEGENNCLLHPMSLHGMEDEDREVLKWK
FDSRLALRHIARERHPEYYKD$

Fig. 95B

2003_CON_F2 nef.OPT

ATGGGCGGCAAGTGGTCCAAGTCCTCCATCGTGGGCTGGCCCACCATCCGCGAGCGCATCCGCCGCACCCCCGTGGCCGCCGAGGGCGTGGG
CGCCGTGTCCCAGGACCTGGACAAGCACGGCGCCATCACCTCCTCCAACACCCGCGCCACCAACGCCGACCTGGCTGGCTGGAGCCCAGG
AGGACGAGGAGGTGGGCTTCCCCGTGCGCCCCCAGGTGCCCCTGCGCCCCATGACCTACAAGGCCGCCTTCGACCTGTCCCACTTCCTGAAG
GAGAAGGGCGGCCTGGAGGGCCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTTCGGCTGGTGTACCACACCCAGGGCTACTTCCCCGA
CTGGCAGAACTACACACCCCCGGCCCCCCGGCACCCGCTACCCGCTGACCTTCGGCTGGTGCTTCAAGCTGGTGCCCGTGGACCCCGAGGAGGTGG
AGAAGGCCAACGAGGGCGAGAACAACTGCCTGCTGCACCCCATGTCCCTGCACGGCATGGAGGACGAGGACCGCGAGGTGCTGAAGTGGAAG
TTCGACTCCCGCCTGGCCCTGCGCCACATCGCCCGCGAGCGCCACCCCGAGTACTACAAGGACTAA

Fig. 96A

45. 2003_CON_G nef.PEP

MGGKWSKSSIVGWPEVRERIRQTPPAAEGVGAVSQDLARHGAITSSNTAANNPDCAWLEAQEEDSEVGFPVRPQVPLRPMTYKGAFDLSFFL
KEKGGLDGLIYSKKRQDILDLWVYNTQGFEPDWQNYTPGPGTRFPLTFGWCFKLVPMDPAEVEEANKGENNSLLHPICQHGMEDEDREVLVW
RFDSSLARRHIARELHPEYYKDC$

Fig. 96B

2003_CON_G nef.OPT

ATGGGCGGCAAGTGGTCCAAGTCCTCCATCGTGGGCTGGCCCGAGGTGCGCGAGCGCATCCGCCAGACCCCCCCGCCGGCAGAGGGCGTGGG
CGCCGTGTCCCAGGACCTGGCCCGCCACGGCGCCATCACCTCCTCCAACACCGCCGCCAACAACCCCGACTGCGCCTGGCTGGAGGCCCAGG
AGGAGGACTCCGAGGTGGGCTTCCCCGTGCGCCCCCAGGTGCCCCTGCGCCCCATGACCTACAAGGGCGCCTTCGACCTGTCCTTCTTCCTG
AAGGAGAAGGGCGGCCTGGACGGCCTGATCTACTCCAAGAAGCGCCAGGACATCCTGGACCTGTGGGTGTACAACACCCAGGGCTTCTTCCC
CGACTGGCAGAACTACACACCCGGGCCACCCGGCTTCCCCCTGACCTTCGGCTGGTGCTTCAAGCTGGTGCCCATGGACCCCGCCGAGG
TGGAGGAGGCCAACAAGGGCGAGAACAACTCCCTGCTGCACCCCATCTGCCAGCACGGCATGGAGGACGAGGACCGCGAGGTGCTGGTGTGG
CGCTTCGACTCCTCCCTGGCCCGCCGCCACATCGCCCGCGAGCTGCACCCCGAGTACTACAAGGACTGCTAA

Fig. 97A

46. 2003_CON_H nef.PEP

MGGKWSKSSIGGWPAIRERIRRAEPAAEGVGAVSRDLDRRGAVTINNTASTNPDSAWLEAQEEEEVGFPVRPQVPLRPMTYKGAFDLSHFL
KEKGGLEGLIYSKKRQEILDLWVYNTQGYFPDWQNYTPGPGERYPLTFGWCFKLVPVDPQEVEKANEGENNSLLHPICQHGMEDEEREVLMW
KFDSRLAFRHIARELHPEFYKDC$

Fig. 97B

2003_CON_H nef.OPT

ATGGGCGGCAAGTGGTCCAAGTCCTCCATCGGCGGCTGGCCCGCCATCCGCGAGCGCATCCGCCGCGCCGAGCCCGCCGCAGAGGGCGTGGG
CGCCGTGTCCCGCGACCTGGACCGCCGCGGCGCCGTGACCATCAACAACACCGCCTCCACCAACCCCGACTCCGCCTGGCTGGAGGCCCAGG
AGGAGGAGGAGGTGGGCTTCCCCGTGCGCCCCCAGGTGCCCCTGCGCCCCATGACCTACAAGGGCGCCTTCGACCTGTCCCACTTCCTG
AAGGAGAAGGGCGGCCTGGAGGGCCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACAACACCCAGGGCTACTTCCC
CGACTGGCAGAACTACACACCCGGGCCCGGCGAGCGCTACCCCCTGACCTTCGGCTGGTGCTTCAAGCTGGTGCCCGTGGACCCCCAGGAGG
TGGAGAAGGCCAACGAGGGCGAGAACAACTCCCTGCTGCACCCCATCTGCCAGCACGGCATGGAGGACGAGGAGCGCGAGGTGCTGATGTGG
AAGTTCGACTCCCGCCTGGCCTTCCGCCACATCGCCCGCGAGCTGCACCCCGAGTTCTACAAGGACTGCTAA

Fig. 98A

47. 2003_CON_01_AE_nef.PEP

MGGKWSKSSIVGWPQVRERIKQTPPATEGVGAVSQDLDKHGAVTSSNMNNADCVWLRAQEEEEVGFPVRPQVPLRPMTYKGAFDLSFFLKEK
GGLDGLIYSKKRQEILDLWVYNTQGFFPDWQNYTPGPGIRYPLCFGWCFKLVPVDPREVEEDNKGENNCLLHPMSQHGIEDEEREVLMWKFD
SALARKHIARELHPEYYKDC$

Fig. 98B

2003_CON_01_AE_nef.OPT

ATGGGCGGCAAGTGGTCCAAGTCCTCCATCGTGGGCTGGCCCCAGGTGCGCGAGCGCATCAAGCAGACCCCCCCGCCACCGAGGGCGTGGG
CGCCGTGTCCCAGGACCTGGACAAGCACGGCGCCGTGACCTCCTCCAACATGAACAACGCCGACTGCGTGTGCCTGCGCGCCCAGGAGGAGG
AGGAGGTGGGCTTCCCCGTGCGCCCCCAGGTGCCCCTGCGCCCCATGACCTACAAGGGCGCCTTCGACCTGTCCTTCTTCCTGAAGGAGAAG
GGCGGCCTGGACGGCCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACAACACCCAGGGCTTCTTCCCCGACTGGCA
GAACTACACACCCCCGGCCCCGGCATCCGCTACCCCCTGTGCTTCGGCTGGTGCTTCAAGCTGGTGCCCGTGGACCCCCGAGGTGGAGGAGG
ACAACAAGGGCGAGAACAACTGCCTGCTGCACCCCATGTCCCAGCACGGCATCGAGGACGAGGAGCGCGAGGTGCTGATGTGGAAGTTCGAC
TCCGCCCTGGCCCGCAAGCACATCGCCCGCGAGCTGCACCCCGAGTACTACAAGGACTGCTAA

Fig. 99A

48. 2003_CON_03_AE_nef.PEP

MGGKWSKSSIVGWPQVRERIRRAPAPAARGVGPVSQDLDKYGAVTSSNTAANNADCAWLEAQKEEEVGFPVRPQVPLRPMTYKGAFDLSHFL
KEKGGLDGLIYSKKRQEILDLWVYHTQGYFPDWQNYTPGPGIRFPLTFGWCYKLVPVDPDEVEEATEGENNSLLHPICQHGMDDEEKEVLMW
KFDSRLALTHRARELHPEFYKDC$

Fig. 99B

2003_CON_03_AE_nef.OPT

ATGGGCGGCAAGTGGTCCAAGTCCTCCATCGTGGGCTGGCCCCAGGTGCGCGAGCGCATCCGCCGCGCCCCAGCTCCAGCTGCGGCTGGCTGGCCCCCCGCGCGCCCGCGGCGT
GGGCCCCGTGTCCCAGGACCTGGACAAGTACGGCGCCGTGACCTCCTCCAACACCGCCGCCAACAACGCCGACTGCGCCTGGCTGGAGGCCC
AGAAGGAGGAGGAGGTGGGCTTCCCCGTGCGCCCCCAGGTGCCCCTGCGCCCCATGACCTACAAGGGCGCCTTCGACCTGTCCCACTTCCTG
AAGGAGAAGGGCGGCCTGGACGGCCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACCACACCCAGGGCTACTTCCCC
GACTGGCAGAACTACACACCCGGCCCCGGCATCCGCTTCCCCCTGACCTTCGGCTGGTGCTACAAGCTGGTGCCCGTGGACCCCGACGAGG
TGGAGGAGGCCACCGAGGGCGAGAACAACTCCCTGCTGCACCCCATCTGCCAGCACGGCATGGACGACGAGGAGAAGGAGGTGCTGATGTGG
AAGTTCGACTCCCGCCTGGCCCTGACCCACCGCGCCCGCGAGCTGCACCCCGAGTTCTACAAGGACTGCTAA

Fig. 100A

49. 2003_CON_04_CFX_nef.PEP

MGGKWSKSSIVGWPAIRERMRQRGPAQAEPAAAGVGAVSQDLDKHGAITSSNTAATNPDKAWLEAQEEEEVGFPVRPQVPLRPMTFKAALD
LSHFLKEKGGLDGLIYSKKRQEILDLWVYNTQGYFPDWQNYTPGPGERFPLCFGWCFKLVPVDPQEVEEATEGENNCLLHPISQHGMEDEER
EVLKWKFDSRLAYKHIARELHPEFYKDC$

Fig. 100B

2003_CON_04_CFX_nef.OPT

ATGGGCGGCAAGTGGTCCAAGTCCTCCATCGTGGGCTGGCCCGCCATCCGCGAGCGCATGCGCCAGCGCGGCCCCGC
CGCCCGCCGCCGCTGGGCGGCGTGGGCGCCGTGTCCCAGGACCTGGACAAGCACGGCGCCATCACCTCCTCCAACAC
GGCTGGAGGCCCAGGAGGAGGAGGAGGGGCGGCTGGGCTTCCCCGTGCCCCGTCCCCAGGTGCCCCTGCGCCCCATGACCTTCAAGGCCGCCCTGGAC
CTGTCCCACTTCCTGAAGGAGAAGGGCGGCCTGGACGGCCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACAACAC
CCAGGGCTACTTCCCCGACTGGCAGAACTACACCCCCGGCCCCGGCGAGCGCTTCCCCCTGTGCTTCGGCTGGTGCTTCAAGCTGGTGCCCG
TGGACCCCCAGGAGGTGGAGGAGGCCACCGAGGGCGAGAACAACTGCCTGCTGCACCCCATCTCCCAGCACGGCATGGAGGACGAGGAGCGC
GAGGTGCTGAAGTGGAAGTTCGACTCCCGCCTGGCCTACAAGCACATCGCCCGCGAGCTCCACCCCGAGTTCTACAAGGACTGCTAA

Fig. 101A

50. 2003_CON_06_CFX_nef.PEP

MGGKWSKSSIVGWPQVRERMRNPPTEGAAEGVGAVSQDLDKHGAITSSNTATTNAACAWLEAQTEDEVGFPVRPQVPLRPMTYKGAFDLSFF
LKEKGGLDGLIYSKKRQEILDLWVYHTQGFFPDWQNYTPGPGIRYPLTFGWCYKLVPVDPKEVEEDTKGENNCLLHPMCQHGVEDEEREVLM
WKFDSSLARRHIAREMHPEFYKDC$

Fig. 101B

2003_CON_06_CFX_nef.OPT

ATGGGCGGCAAGTGGTCCAAGTCCTCCATCGTGGGCTGGCCCCAGGTGCGCGAGCGCATGCGCAACCCCCCCACCGAGGGCGCCGCCGAGGG
CGTGGGCGCCGTGTCCCAGGACCTGGACAAGCACGGCGCCATCACCTCCTCCAACACGGCCACCACGAACGCCGCCTGCGCCTGGCTGGAGG
CCCAGACCGAGGACGAGGTGGGCTTCCCCGTGCGCCCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACCACACCCAGGGCTTCTT
CCCCGACTGGCAGAACTACACCCCCGGCCCCGGCATCCGCTACCCCCTGACCTTCGGCTGGTGCTACAAGCTGGTGCCCGTGGACCCCAAGG
AGGTGGAGGAGGACACCAAGGGCGAGAACAACTGCCTGCTGCACCCCATGTGCCAGCACGGCGTGGAGGACGAGGAGCGCGAGGTGCTGATG
TGGAAGTTCGACTCCTCCCTGGCCCGCCGCCACATCGCCCGCGAGATGCACCCCGAGTTCTACAAGGACTGCTAA

Fig. 104A

53. 2003_CON_11_CFX nef.PEP

MGGKWSKSSIVGWPEIRERLRRTPPTAAAEGVGAVSKDLEKHGAVTSSNTAQTNAACAWLEAQEEEVGFPVRPQVPLRPMTYKGAFDLGFF
LKEKGGLDGLIYSKKRQEILDLWVYHTQGYFPDWQNYTPGPGIRYPLCFGWCFKLVPVEPREVEEANEGENNCLLHPMSQHGMDDEEREVLM
WKFDSSLARRHIARELHPDFYKDC$

Fig. 104B

2003_CON_11_CFX nef.OPT

ATGGGCGGCAAGTGGTCCAAGTCCTCCATCGTGGGCTGGCCCGAGATCCGCGAGCGCCTGCGCCGCACCCCCACCGCCGCCGCCGAGGG
CGTGGGCGCCGTGTCCAAGGACCTGGAGAAGCACGGCGCCGTGACCTCCTCCAACACCGCCCAGACCAACGCCGCCTGCGCCTGGCTGGAGG
CCCAGGAGGAGGAGGAGGTGGGCTTCCCCGTGCGCCCCCAGGTGCCCCTGCGCCCCATGACCTACAAGGGCGCCTTCGACCTGGGCTTCTTC
CTGAAGGAGAAGGGCGGCCTGGACGGCCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACCACACCCAGGGCTACTT
CCCCGACTGGCAGAACTACACCCCCGGCATCCGCTACCCCCTGTGCTTCGGCTGGTGCTTCAAGCTGGTGCCCGTGGAGCCCCGCGAGTG
AGGTGGAGGAGGCCAACGAGGGCGAGAACAACTGCCTGCTGCACCCCATGTCCCAGCACGGCATGGACGACGAGGAGCGCGAGGTGCTGATG
TGGAAGTTCGACTCCTCCCTGGCCCGCCGCCACATCGCCCGCGAGCTGCACCCCGACTTCTACAAGGACTGCTAA

Fig. 105A

54. 2003_CON_12_BF nef.PEP

MGGKWSKSSIVGWPDIRERMRRAPPAAEGVGAVSQDLENRGAITSSNTRANNPDLAWLEAQEEEVGFPVRPQVPLRPMTYKGALDLSHFLK
EKGGLEGLIYSKKRQEILDLWVYHTQGYFPDWQNYTPGPGIRYPLTFGWCFKLVPVDPEEVEKANEGENNCLLHPMSQHGMEDEDREVLMWK
FDSRLALRHIAREKHPEFYQDC$

Fig. 105B

2003_CON_12_BF nef.OPT

ATGGGCGGCAAGTGGTCCAAGTCCTCCATCGTGGGCTGGCCCGACATCCGCGAGCGCATGCGCCGCGCCCCCGCCGCCGAGGGCGTGGGG
CGCCGTGTCCCAGGACCTGGAGAACCGCGGCGCCATCACCTCCTCCAACACCCGCGCCAACAACCCCGACCTGGCCTGGCTGGAGGCCCAGG
AGGAGGAGGAGGTGGGCTTCCCCGTGCGCCCCCAGGTGCCCCTGCGCCCCATGACCTACAAGGGCGCCCTGGACCTGTCCCACTTCCTGAAG
GAGAAGGGCGGCCTGGAGGGCCTGATCTACTCCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACCACACCCAGGGCTACTTCCCCGA
CTGGCAGAACTACACCCCCGGCCCCGGCATCCGCTACCCCCTGACCTTCGGCTGGTGCTTCAAGCTGGTGCCCGTGGACCCCGAGGAGGTGG
AGAAGGCCAACGAGGGCGAGAACAACTGCCTGCTGCACCCCATGTCCCAGCACGGCATGGAGGACGAGGACCGCGAGGTGCTGATGTGGAAG
TTCGACTCCCGCCTGGCCCTGCGCCACATCGCCCGCGAGAAGCACCCCGAGTTCTACCAGGACTGCTAA

Fig. 106A

55. 2003_CON_14_BG_nef.PEP

MGGKWSKCSIVGWPEVRERIRRTPPAAVGVGAVSQDLAKHGAITSSNTAANNPDCAWLEAQEEDSEVGFPVRPQVPLRPMTYKGAFDLSFFL
KEKGGLDGLIYSKQRQDILDLWVYNTQGFFPDWQNYTPGPGTRYPLTFGWCFKLEPVDPAEVEEATKGENNSLLHPICQHGMEDADNEVLIW
RFDSSLARRHIARELHPDFYKDC$

Fig. 106B

2003_CON_14_BG_nef.OPT

ATGGGCGGCAAGTGGTCCAAGTGCTCCATCGTGGGCTGGCCCGAGGTGCGCGAGCGCATCCGCCGCACCCCCCCGGCCGCCGTGGGCGTGGG
CGCCGTGTCCCAGGACCTGGCCAAGCACGGCGCCATCACCTCCTCCAACACCGCCGCCAACAACCCCGACTGCGCCTGGCTGGAGGCCCAGG
AGGAGGACTCCGAGGTGGGCTTCCCCGTGCGCCCCCAGGTGCCCCTGCGCCCCATGACCTACAAGGGCGCCTTCGACCTGTCCTTCTTCCTG
AAGGAGAAGGGCGGCCTGGACGGCCTGATCTACTCCAAGCAGCGCCAGGACATCCTGGACCTGTGGGTGTACAACACCCAGGGCTTCTTCCC
CGACTGGCAGAACTACACCCCGGGCCCGGGCACCCGCTACCCCCTGACCTTCGGCTGGTGCTTCAAGCTGGAGCCCGTGGACCCCGCCGAGG
TGGAGGAGGCCACCAAGGGCGAGAACAACTCCCTGCTGCACCCCATCTGCCAGCACGGCATGGAGGACGCCGACAACGAGGTGCTGATCTGG
CGCTTCGACTCCTCCCTGGCCCGCCGCCACATCGCCCGCGAGCTGCACCCCGACTTCTACAAGGACTGCTAA

Fig. 107A

61. 2003_2003_CON_S_pol.PEP

FFRENLAFQQGEAREFSSEQTRANSPTSRELRVRGGDNPLSEAGAERQGTVSLSFPQITLWQRPLVTVKIGGQLKEALLDTGADDTVLEEIN
LPGKWKPKMIGGIGGFIKVRQYDQILIEICGKKAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEK
IKALTEICTEMEKEGKISKIGPENPYNTPIFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDE
DFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRTQNPEIVIYQYMDDLYVGSDLEIGQHRTKIEELREHLLRWGF
TTPDKKHQKEPPFLWMGYELHPDKWTVQPIQLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVKQLCKLLRGAKALTDIVPLTEEAELELAEN
REILKEPVHGVYYDPSKDLIAEIQKQGQDQWTYQIYQEPFKNLKTGKYAKMRSAHTNDVKQLTEAVQKIATESIVIWGKTPKFRLPIQKETW
ETWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIVGAETFYVDGAANRETKLGKAGYVTDRGRQKVVSLTETTNQKTELQAIHLALQDSG
SEVNIVTDSQYALGIIQAQPDKSESELVNQIIEQLIKKEKVYLSWVPAHKGIGGNEQVDKLVSTGIRKVLFLDGIDKAQEEHEKYHSNWRAM
ASDENLPPIVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKVIH
TDNGSNFTSAAVKAACWWAGIQQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIAT
DIQTKELQKQITKIQNFRVYYRDSRDPIWKGPAKLLWKGEGAVVIQDNSEIKVVPRRKAKIIRDYGKQMAGDDCVAGRQDED$

Fig. 107B

2003_CON_S pol.OPT

TTCTTCCGCGAGAACCTGGCCTTCCAGCAGGGCGAGGCCCGCGAGTTCTCCTCCGAGCAGAGACCCGGCCAACTCCCCCACCTCCCCGCGAGCTGCGCGTGCG
CGGCGGCGACAACCCCCTGTCCGAGGCGGGCGGCCGAGGCGCCGAGGCGCCCTGCTGGACACCGGCCGCGACGACACCGGTCCTTCCCCCAGATCACCCTGTGGCAGCGCCCCTGGTGACCG
TGAAGATCGGCGGCGCCATCGGCGGCTTCAAGGTGCGCCAGTACGACCAGATCTGATCGAGATCTGATCGAGAAGGCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCG
ATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCTGATCGAGATCTGATCGAGATCTCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCG
CCCCGTGAACATCATCGGCCGCAACATGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGAGATCTGCACCGAGATCTGCACCGAGATGGAGAAGGAGGCCAAGATCTCC
AAGATCGGCCCCGAGAACCCCTACAACACCCCATCTTCGCCATCAAGAAGAAGGACTCCACCAAGTGGCCCAAGCTGGTGGACTTCCGCGAGCTGAACAA
GCGCACCCCAGGACTTCTGGGAGGTGCAGCTGGCAGCTTCCGCAAGTACACCGCCTTCACCATGACCAAGATCCTCCAGTCGTGATCGAGTGATCGATCTACCAGTA
TCTCCGTGCCCCTGGACGAGGGCTCCCCCGCCATCTTCCAGTCCTCCATGACCAACCGGCCAGCAGATCGGAGATCGGAGAGCTGCACCACCCC
CCGACAAGAAGACACCAGAGAAGGAGCCCCCTTCCTGTGGACGTGCGCTACGAGCTGCAGCAGTGACCCCGGCATCAAGGTGACCAGTCGTGCACGGCGTGT
TCCTGGACCGTGAACGACATCCAGAAGCTGGTGGCCCCTGACCGTGCCCCTGACCGAGGAGGCCGAGCTGGACCCTGACCGAGAACCGGAGATCTACCAGATCTACCCGGCATCAAGGTGACCAGTCGTGCGGCGTGCGCG
CGGCGCCAAGGGCCTGTACGTGGGCCTGACCGAGGAGATCGGAGACCAGATCCAGAAGAACCGGAGATCCTGATCGTGGGCAA
ACTACGACGCCAAGAATGCGCTCCGCCCTCCAAGGACCTGATCGCGCGCCCACCACCAGATCGCCGAGAAGATCGCCGAGTCCATCGTCGTGGGCGAACA
GGCAAGTACGCCAAGAATGCGCCCTGCCCATCCAGAAGGAGACCTGGTGGACCAGTACTGGACCCGCGCCCAACCGCGCCAAGCTGGTCCCCGCGCCCAGTACTGGACCCGCGCCCAACCGCGCCAAGCTG
GACCCCCAGTTCCGCTGCCCATCCAGAAGGAGACCTGGTGGACCAGTACTGGACCGAGTTCGTGAACA
CCCCCCCCTGGTGCAAGTTCCGCTGTGGTACCAGTCGCGGCGGCCCATCTCGTGGGCGGCGCCAACCGCGAGACCAAGCTG
GGCAAGGCCGGCTACGTGACCGAGGTGGTCCCTGACCGCGCCCAGTACTGACCCTGGGCCCCGCGGAGATCTCCGAGCTGGTGAACC
GCAGGACTCCGGCTCCGAGGTGAACATCGTGACCGAGGAGGTGTACCGTGCCTGTCCTGGGCGCCGCGCAGGAGCCCAGAGGTGGACAAGCTGGTGTCC
AGATCATCGAGCAGCTGATCAAGAAGGACGAGAGGGCCATCGAGGGCCCCGGCATCT
CCTGCCCCCATCGTGGCCCAAGGAGATCGTGGCCTCCTGCGCCGAGTCGTGCCCGTGAAGGTGATCCCCGCGCGGAGACCGGC
GGCAGCTGGACTGCGCGCCACCCTGGAGGGGCCAAGATCATCCTGTGCCCTGGTGCCCGTGCCCGTACCGGCGTGGTGGGCGTACCCTGGAGGTGGTCCCCGCGCTCAGAGGC
CAGGACCTGCTGCAGGTGCGCGGCCAGCAGGAGTTCGGCATCCAGCGGAGTTCGGCATCCAGCGGCCCGTGGGGAGTCGAAGAGATCA
TCGGCAGGTGCGCGGCGACCAGCGGGCCGAGCGCCATCCAGCGGCCCGTGGGCCGTGTTCATCCAACTTCAAGCGCGGCTACTCC
GCCGGCAGCGAGGGCCATCATCGACGACATCATCGCGCCCAAGCTGCTGTGGAAGGGCGTGGGGGAAGACCTGAGATCCAGAGGCCGA
CTCCCCGCGACCCCATCTGGAAGGGCGTGGTACGGCGTGGGCGCCCAAGGACCACGGGTGGTGGGCCAAGGTGGTGCCCGCC
GCAAGGCCAAGATCATCCGCGACTACGGCGGCGGCGTGCGGCTGGGGCGGCGGCGAGATGGCCGCGGCCAGCAGCAAGGGCCAAGGCCAAGTACGGCGGCGGCGTGCGGCTGGGGGCGCCGCGATGGCCGCGGCCAGCAGCAAGGGCCAAGGCCAAGTACGGCGGCGGCGTGCGGCTGGGGGCGCCGCGAGCAGGACGAGGACTAA

Fig. 108A

62 2003_M GROUP anc pol.PEP

FFRENLAFQQGEAREFSSEQTRANSPTSRELRVRGGDNPLSEAGAERQGTVSFSFPQITLWQRPLVTIKIGGQLREALLDTGADDTVLEEIN
LPGKWKPKMIGGIGGFIKVRQYDQILIEICGKKAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEK
IKALTEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDE
TTPDKKHQKEPPFLWMGYELHPDKWTVQPIQLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVKQLCKLLRGAKALTDIVPLTEEAELELAEN
REILKEPVHGVYYDPSKDLIAEIQKQGQDQWTYQIYQEPFKNLKTGKYAKMRSAHTNDVKQLTEAVQKIATESIVIWGKTPKFRLPIQKETW
ETWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIVGAETFYVDGAANRETKLGKAGYVTDRGRQKVVSLTETTNQKTELQAIHLALQDSG
SEVNIVTDSQYALGIIQAQPDKSESELVNQIIEQLIKKEKVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGIDKAQEEHEKYHSNWRAM
ASDFNLPPVVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKVILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKVIH
TDNGSNFTSAAVKAACWAGIQQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIAT
DIQTKELQKQITKIQNFRVYYRDSRDPIWKGPAKLLWKGEGAVVIQDNSEIKVVPRRKAKIIRDYGKQMAGDDCVAGRQDED$

Fig. 109A

63. 2003_CON_A1 pol.PEP

FFRENLAFQQGEARKFSSEQTGANSPTSRDLWDGGRDSLPSEAGAERQGTGPTFSFPQITLWQRPLVTVRIGGQLKEALLDTGADDTVLEDI
NLPGKWKPKMIGGIGGFIKVKQYDQILIEICGKKAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEE
KIKALTEICTEMEKEGKISKIGPENPYNTPIFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLD
FTTPDKKHQKEPPFLWMGYELHPDKWTVQPIELPEKESWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRGAKALTDIVTLTEEAELELAE
NREILKDPVHGVYYDPSKDLIAEIQKQGQDQWTYQIYQEPFKNLKTGKYARKRSAHTNDVKQLAEVVQKVVMESIVIWGKTPKFKLPIQKET
WETWWMDYWQATWIPEWEFVNTPPLVKLWYQLEKDPIVGAETFYVDGAANRETKLGKAGYVTDRGRQKVVSLTETTNQKTELHAIHLALQDS
GSEVNIVTDSQYALGIIQAQPDRSESELVNQIIEKLIGKDKVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGIDKAQEEHERYHSNWRA
HTDNGSNFTSAAVKAACWANIQQEFGIPYNPQSQGVVESMNKELKKIIGQVREQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIA
TDIQTKELQKQITKIQNFRVYYRDSRDPIWKGPAKLLWKGEGAVVIQDNSDIKVVPRRKAKIIRDYGKQMAGDDCVAGRQDED$

Fig. 108B

2003_M.GROUP anc pol.OPT

TTCTTCCGCGAGAACCTGGCCTTCCAGCAGGGCGAGGCCCGCGAGTTCTCCTCCGAGCAGACAGACCCGGCGCCAACTCCCCACCTCCCCGCGAGCTGCGCGCTGCG
CGGCGGCGACAACCCCGTGTCCGAGGCCGGCGCCGAGGCCCGAGCCCTGGACACCGGCGCGCCTGCTGGAGGCACCGTGTCCTTCCCCAGATCACCCTGTGGCAGCCCCTGGTGACCA
TCAAGATCGGCGGCGGCGGCGAGGCCCAGTCCTGGACAACCGGCGGCGCTGCTGGAGGAGATCAACCTGCTGGAGGATCGGCGGCAAGAGGCCATCGGCGCACCGTGTCGGTGGGCCCAC
ATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCTGATCGGCTGCCACCTGCCCCATCGAGACTTCCCCATCTCCCCGATGGAGGCCGTGCCCGTGAAGCTGAAGCCCG
CCCGGTGAACATCATCGGCCGCAACATGCTGAAGGTGGCCCCAAAGGCCCCAGATCTGCACCGGAGATGGAGAAGGGCAAGATCTCC
GCATGGACGGCCCAGATGGAAGCCAGTGCCCCTGACCGAGGAAGATCAAGGCCCTGACCGAGATCTGCACCCAAGTGGCCGGAGCTGAACAA
AAGATCGGCCCGCGAGAACCCCTACACACCCCCGTGTTCGCCATCAAGAAGAAGGACTCCACCAAGTGGCGCAAGCTGGTGGACCGTGGGGACGGCCTACT
GCGCACCCAGGACTTCTGGGAGGTGCAGCTGGCATCCCCACCCGCCTTCACCAGATCCTCCAGTCCTTCCATGACCAAGATCCTGAGAGATCGTGATCTACCAGTA
TCTCCGTGCCCCTGGACGAGGAGGACTTCCGCAAGTACACACCGCCTTCCAGATCGGCCAAGAACCCGAGAGAATCCGCACCAAGGACGTGCGCGAGAGCTGTCGCG
CCCCAGGGCTGGAAGGGCTCCCCCGCCATCTTCCAGTCTTCCAGGCCAGCTGGACCTGGCCAAGAGTCCTGCGCTGCTCGGGGCTTCACCACCC
CATGGACGACCTGTACGTGGGCTTCCGAGATCGGCCTGGACATCGGCCAGAGCTGCGCGCCAAGTGGGCTACGAGAGCTGTGTCAAGCAGCTGTCGCG
CCGACAAGAAGCACCAGAGAGAGCAGCCCCCCGTGACGCTACCCCGAGAACCGCGAGATCTACCCGGCCGAGAACCGCGAGATCTACCAGGAGAGCCCTTCAAGAACCTGAAGACC
TCCTGGACCGTGAACGACATCCAGAGACCGCCGAGCTGAACTGGGCGCTGAAGCTGATGGCGGCGGAGTTCGTGAACA
CGGCGCCAAGGGCCCTGACCGAGAATCGGCTGCCTCCGCCATCCAGAACGACGTGGGCGCCGAGAACCGCGAGACCAAGCTG
ACTACGACCCCTTCCAAGGACCTGATCGCCGCCCCACACCAACGACCTGGGGCGCCGTGACTGGTGGGAGTTCGTGAACA
GGCAAGTACGACGCCAGAGATGCGCCCGCGCCATCCAGAAGGAGAGCCCCAGAGAAGGTGGTGTCCCTGACCGAGACCGAGCTGCAGCCACCTGGCCCT
GCCAGGACTCCGGCTCCGAGGTGAACATCGTGACCGAGCTCCCCAGTACGCCCTGGGCCATCATCCCGAGTCCGAGTCCGTGGTGAACC
AGATCATCGAGAACGCAGCTGATCAAGAGGAGAAGGTGTACCGTGTCCTGGGCGCCGCCCACAGGGCATCGGCGCAACGAGCAGTGACAAGCTGGTGTCC
TCCGGCATCGGCAAGGTGCTGTTCCTGGACGGCATCGACAAGGCCAGGAGCACGAGGAAGTACCACTCCAACTGGCGCGCCTCCGCCCCCGGCATCT
CCTGCCCCTGGACTGGACCACCGTGGTGGCCAAGGAGATCTCATCGTGGCCCGCTGGCCGCGTGAAGGTGATCCTGGCCGTGGCCCGCGCCATCCGTGAAGGC
CAGGAGACCGCCGCCTACTTCATCCTGAAGCTGGCCGGCAGGAGCCCAGTGCTGGCCTGCTCCCCCGCGCCGAGACCGGC
CGCCTGCTGTGGGCGCGCCAAGGCCCTACACGCGCACCTCCCTACAACAACCCCCAGTCCCAGGCGTGGTGAGTCCATGAACAAGGCGGCTACTCC
TCGGCCAGGTGCGCGACCAGGCCGACTGAAGACCCGAGATGGCCGTGTTCATCCACAACTTCAAGCGCAAGGGCGGCATCGGCGGCTACTACGGCGA
GCCGGCGAGCGGCATCATCGACATGGCCCCCGCCCAAGCTGCTGTGTGGAAGGGCGGCGCCGTGTGGTCGGCGGCATCAAGGTGTGTGCCCCCGCC
CTCCCCCCCCGGCCAAGAATCATCCGCGGCAAGGCCAAGAGATCATCGGCGACTACGGCAAGCAGATGGCCGGCAAGGCCAAGAGCAGCGAGGCGAGAGGGACTAA

Fig. 109B

2003_CON_A1 pol.OPT

```
TTCTTCCGCGAGAACCTGGCCTTCCAGCAGGGCGAGGCCCGCGGCCGCCAACTCCCCACCTCCCGCGACCTGTGGGACGG
CGGCCGCGGACTCCCTGCCCTCCGAGGCCGCGGCGCCGAGCGCCCAGGGCACCGGCACCCTGTGGCAGCCGCCCCTGGTGA
CCGTGCGCATCGGCGGCGGCCAGCTGAAGGAGGCCCTGCTGGACACCGGCGCCGACGACACCGTGCTGGAGGAGCCCAAG
ATGATCGGCGGCATCGGCGGCTTCATCAGCGTGAAGCCGCGCAACATGCTGGCCGCAGATCGGCTGCGAGATCGGGGCCC
CACCCCCGTGAACATCATCGGCGGCCGCAACATGCTGACCCAGATCGGCTGCACCCTGAACTTCCCCATCCCCCGTGAAGC
CCGGCATGGACGCGCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGAGATCTGCACCGAGA
TCCAAGATCGGCACCACCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCCGGCCTGAAGAAGAAGAAGAGATC
CAAGCGCACCGCCAGGACTTCTGGGAGGTCCTTCCGCCAGGTGCAGCTGGGCATCCGCGCCATCCCCCACCCCGCCGGCCT
ACTTCTCCGTGCCCGTGGACACTTCTGGACGAGTCTTCCGCCAGGTCCTTCCATGATGACCAGATCCTGGAGATCGAGGTG
CTGCCCCCCAGGGGCTGGAAGGGCTCCCCCGGCCTCCGCGACAGCTGGCCAAGTACACCGCCTTCCAGATCCTCCAAGATCCTG
GTACATGACGACCTGTACGTGGCTCCGACTGGGCCAAGCTGTGTGGGCAAGTCTTCCGCCATCTTCCGACCTGTCCTGGGCTTCACCA
CCCCCGACAAGAAGCACCAGAAGGAGCTCCCCGGCCTTCCTGTGGATGGGCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGG
GAGTCCTGGACGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCGGGCATCAAGGTGAAGCTGCT
GCGCGGCGCCAAGGCCCTGACCGATGTGGTCCCCCTGACCGAGGAGGCAGGGCCGAGAGCAGGGGCCGAGAGTCCTACCAGGAAGCTTCAAGAACCTGAAG
TGTACTACGACCCCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGGTGGTCAGAAGTGGTGATGGGACTGGGGG
ACCGGCAAGTACGCCCGGGCCAAGCTGCCCCATCCAGAAGGAGACCTGGGAGATGTGGGCGCGCCCCATCTCGTGGGAGTTCGTGA
CAAGAACCCCCCGGCGCAAGTTCAAGCTTGGTGAAGCTGTGTGATGCCGACCGGCCAGCCCCAGGCCGCGCCCAACCCGAGACCAAG
CTGGGCAAGGCCGGCTACGTGGCTGGCCGCGCCCCATCGTGGGAGTTCGTGTACCGTGACCGGCATCAGCCGCGCCATCCACCTGGC
CCTGCAGGACTCCGGAGGTGAACATCGTGACCGACTCCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCGGACAAGGTGATCAGCAAGGTGGTG
ACCAGATCATCGAGAAGCTGATCGGCAAGGTGCTGTTCCTGGACGGCATCGACAAGGCCCAGGAGCACGAGCGCTACCACTCCAACTGGCGTGGTG
TCCTTCCGGCATCGCCCCCCATCGTGGCCAAGGAGATCGTGGCCAGCTGCGACAAGTGCCAGCTGAAGGGCGAGGCCATGGCGCGCTCCGACTT
CAACCTGCCGCCTGGACTGCACCCACGTCGTGCCGCAGATCCTGGTGGCCGTGCACGTGGCCAGCGGCTACATCGAGGCCGAGGTGATCCCCGCA
TCTGGCAGCTGGACTGCACCCACCTTCCTGCTGAAGCTGGCCGGCCGCTGGCCCGTGACGTGGTGCACACCGACAACGGCTCCAACTTCACCTCCGCCGTGAA
GGCCAGGAGGAGCGCCTACTTCGGCCTGGCCGCCATCCAGGCCCAGGAGTTCGGCATCCCCTACAACCCCCAGAGTCCAGGGCGGGCATCATCGAGTCC
TCATCGGCCAGGTGCGCGACCAGGCCGAGCACCTGAAGACCGCCGTGTTCATCCAGAACTTCAAGCGCAAGGGGCGGCATCGGCGGCTAC
TCCGCCGCCGAGCGCATCGTGGACATCATCGCCACCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCATCAAGATCCAGAACTTCCGCGTGTACTACCG
CGACTCCCGCGACCCGCTGTGGAAGGGCCCGCTGCGCGTGGACGACCGGGGGCCGCGCGGCCGAGATGGCCGGCGACGCTGCGTCAAGGTGGTGCCCC
GCCGCCAAGGGCGCCGACTACGGCCAAGCAGCAGATGGCCGGCGACGACTGCGTCAAGGTGGTGCCCCGCCAGGGCGCCGACTGCGTCAAGGTGGTGCCCC
GCCGCCAAGGGCGCCGACTACGGCCAAGCAGCAGATGGGCCGGGACGACTGCGTCAAGGTGGTGCCCCGCCAGGGCGCCGACGAGGACTAA
```

Fig. 109C

64. 2003_A1.anc pol.PEP

```
FFRENLAFQQGEARKFSSEQTRANSPTSRELWDGGRDSLLSEAGAERQGTVPSFSFPQITLWQRPLVTVKIGGQLKEALLDTGADDTVLEDI
NLPGKWKPKMIGGIGGFIKVRQYDQILIEICGKKAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEE
KIKALTEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLD
ESFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRSKNPEIVIYQYMDDLYVGSDLEIGQHRAKIELRAHLLSWG
FTTPDKKHQKEPPFLWMGYELHPDKWTVQPIKLPEKDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRGAKALTDIVTLTEEAELELAE
NREILKDPVHGVYYDPSKDLVAEIQKQGQDQWTYQIYQEPFKNLKTGKYAKKRSAHTNDVKQLTEVVQKVATESIVIWGKTPKFRLPIQKET
WETWWMEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIAGAETFYVDGAANRETKLGKAGYVTDRGRQKVVSLTETTNQKTELHAIHLALQDS
GSEVNIVTDSQYALGIIQAQPDRSESELVNQIIEKLIEKEKVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGIDKAQEEHEKYHSNWRA
MASDFNLPPIVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKVILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKVV
HTDNGSNFTSAAVKAACWWANIQQEFGIPYNPQSQGVVESMNKELKKIIGQVREQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIA
TDIQTKELQKQITKIQNFRVYYRDSRDPIWKGPAKLLWKGEGAVVIQDNSDIKVVPRRKAKIIRDYGKQMAGDDCVAGRQDED$
```

Fig. 109D

2003_A1.anc pol.OPT

```
TTCTTCCGCGAGAACCTGGCCTTCCAGCAGGGCGAGGCCCGCCAAGTTCTTCCTCCGAGCAGACCCGCGAGCTGTGGGACGG
CGGCCGCGACTCCCTGCTGTCCGAGGCCGGCGGCCGAGCGGCCACCGTGCCCTCCTTCCTCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGA
CCGTGAAGATCGGCGGCGCCAGCTGAAGGAGGCCCTGCTGGACACCGGCGCCGACGACCAGAGATCCTGGAGGACATCAACCTGCCCGGCAAGTGGAAGCCCAAG
ATGATCGGGGCATCGGCGGCTTCATCGAAGGTGCGCCAGTACGACCAGAGATCCTGGGCTGACCCTGACCCGCGACCGTGCCCGTGGTGGCCC
CACCCCCGTGAACATCATCGGCGCCGAACATGCTGGCCCCTGACCCAGGAGGAAGATCAAGGCCCTGACCGAGATCTGCACCGAGATGGAGAAGGGCAAGATC
CGGCATGGACGGCCCCGGAGAACCCCTACAACACCCCGTGTTCGCCATCAAGAAGAAGGACTCCACCAAGTGGCGCAAGCTGGTGGTGCTGCGCGACTGAA
TCCAAGATCGGCGGCACCAGGACTTCTGGGAGGTGCAGCTGGCCATCCCCACCCCCCGTGTTCGCCATCAAGAAGAAGGACTCCACCAAGTGGCGCAAGCTGGTGGTGCTGCGCGACTGAA
CAAGCGGCACCCAGGACTTCTGGGAGGTGCAGCTGGCCATCCCCACCCCCGTGTTCGCCATCAAGAAGAAGGACTCCACCAAGTGGGCGTGGGCGACGCCT
ACTTCTCCGTGCCCCTGGACAGGGCTCCTCCCGCCAATCTTCCAGTCCTCCATGACCAAGATCCTGGAGCCCTTCCGCTCCAAGAACCCCGCCCCTGCCCCGCCACCAGTACAACGTG
CTGCCCCCAGGGCTGGAAGGGCTCCCCGCTGGGCTCCGACCTGTACGTGGGCTACACGGACTTCCTGCTGCCTGTGTCCTGGGCTTCACCA
GTACATGGACGACCTGTACGTGGGCTCCGACTGTGGCATCGGCCAGCTGCTACGAGCTGGCCTCCGAGCTGTTCCTGTGTCCTGGGCTTCACCA
CCCCCGACAAGAAGCACCAGAAGAGCCCCCTTCCTGTGTGATGGGCTACGAGCTGGCCTCCGAGCTGCTGCCCAGAAGTGGAAGCTGCCCGAGAAG
GACTCCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAGCCCCGAGATCTGAACTGGGCTCCGAGCTGAACGCTGGGCTGCAAGCTGCT
GCGCGGCGCCAAGGCCCTGACCGAGGTGGTCGGCGGCTCCGGAGATCCAGAGAAGCAGGGGCCCAGGGCGAGGAACCGGCCCCTTCAAGAACCTGAAG
TGTACTACGACCCCCTCCAAGACCGTGGTGGCCCCCACACCAACGACGCTGAAGAGCTGAGGTGGTGCAGAGCAGCTGCCGAGGTTCATGCTGTGATCGTGGGG
ACCGGCAAGTACGCCGCCAAGAAGCGCTCGCCCTGCCCATCCCAGAAGGAGAACCTGGACCGCGCCCATCTCGGATCGCCGAGTCCATCGTGATCTGGGG
CAAGAGACCCCCAAGTTCCGCTGCCTGCCCATCCCAGAAGGAGAACCTGGACCGCGCCCATCTCGGATCGCCGAGTCCATCGTGATCTGGGG
ACACCCCCCCAGTTCCGCTGGAAGCTGTGGTACCGACCGGCGGCAACCGCGAGAAGGACCCCATCCCCGGCATGATCCCCGAGTGGGAGTTCGTGA
CTGGGCAAGGGCTACGTGACCGGCGGCAACATCGTGACCGACTCCCAGTACGACCCCTGGGCTGTCCCCAGTACGACCGGCGGCAACCGCGAGAAGGACCCCATCCCCGGCATGATCCCCGAGTGGGAGTTCGTGA
CCTGCAGGAGGACCGGCTACGTGACCGACTCCCAGTACGACCCCTGGGCTGTCCCCAGTACGACCGGCAGTGGACAAGCTGGTG
ACCAGATCATCGGAGAAGCTGATCGAGAAGGAGAAGGTGTACCTGTCCTGGGTGCCCGCCAACAGAGAAGTACCACTCCAACAAGGAGCTGAAGAAGA
TCCTCCGGCATCGGTGGGCGCGAGCGCATCATCGCCGAGCAGGTTCGGCATCAAGAAGCTGAAGGGCGGCATCGGCGGCTAC
CAACCTGCCCCCCCAGTTCCGCGGCGAGCGCATCATCGCCGAGGTGCGGCATCATCCGCCCAGCATCATCGGCGTGTTCATCCACAACTTCAAGCGCGGCTAC
TCTGGCAGCTGGAAGCTGTGGTACCACCCACCTTCTGCTGCTGAAGCTGTGGTGGCCCGGCTGCCTGGAAGGTGATCCCGCCGCCGAGACC
GGCCAGGAGACCGCGCCTACTTCCTGCTGAAGCTGTGGCCCGCTGGCCCGTGAAGGTGGTGCACACCGGGCTCCAACTTCACCTCCGCCGCCGTGAA
GGCCGCCTGCTGGGCCAACATCGGCCAGCAGAGTTCGGCATCAACACCCCCAGTCCGGTGGTGGAGTCCATGAGCGGCATCATCGGCGGCTAC
TCATCGGCCAGGTGCGCGAGCGCATCATCGACGATCATCGAGGGCCCCGCCGTGTTGGAAGCTGTGGTGCCCGGCATCATCCGCCGTGTTCATCCACAACTTCAAGCGCGGCTAC
TCCGCCGGCGGCATCGGCCCCATCTGTGAAGGGCGACCCCATCTGCCGCCAAGCTGCTGTGGTGGAAGCTGTGGTGCCCGGCTGCCTGGAAGGTGATCCCGCCGCCGAGACC
CGACTCCCGCCAAGGCCAAGATCATCCCGCGACTACGGCAAGCAGCAAGATGGCCGCCGGAGGGGCGAGGGCGAGGACGAGGAGGAC
GCCCGCCAAGGCCAAGATCATCCGCGACTACGGCAAGCAGCAAGATGGCCGCCGGAGGGGCGACTGCTGTGTGCCCC
GCCCGCCAAGGCCAAGATCATCCGCGACTACGGCAAGCAGCAAGATGGCCGCCGGAGGGGCGACTGCTGTGGTGTGCCCC
GCCCGCCAAGGCCAAGATCATCCGCGACTACGGCAAGCAGCAAGATGGCCGCCGGAGGGGCGACGAGGAGGACTAA
```

Fig. 110A

65. 2003_CON_A2 pol.PEP

```
FFRENLAFQQREARKFSSEQNRANSPTSRELRNGGRDNLLSEAGAEEQGTVHSCNFPQITLWQRPLVTVKIEGQLREALLDTGADDTVLEDI
NLPGKWKPKMIGGIGGFIKVRQYDQIAIEICGKRAIGTVLVGPTPVNIIGRNMLVQLGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEE
KIKALTEICKEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLH
EDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRSKNPEMVIYQYMDDLYVGSDLEIGQHRAKIEELRAHLLRWG
FTTPDKKHQKEPPFLWMGYELHPDKWTVQPIKLPEKDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRGTKALTDIVTLTKEAELELEE
NREILKNPVHGVYYDPSKDLIAEIQKQGQDQWTYQIYQEPFKNLKTGKYAKRKSTHTNDVKQLTEAVQKIAIESIVIWGKTPKFRLPIQKET
WETWWTEYWQATWIPEWEFVNTPPLVKLWYQLETEPIAGAETFYVDGAANRETKLGKAGYVTDRGRQKIVSLTETTNQKTELHAIYLALQDS
GLEVNIVTDSQYALGIIQAQPDRSESELVNQIIEKLIEKERVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGIDKAQEEHERYHSNWRA
MAHDENLPPIVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKVILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKVI
HTDNGPNFTSATVKAACWWAGVQQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIA
TDIQTKELQKQIIKIQNERVYRDSRDPIWKGPAKLLWKGEGAVVIQDNSDIKVVPRRKAKIIRDYGKQMAGDDCVAGRQDED$
```

Fig. 111A

66. 2003_CON_B pol.PEP

```
FFREDLAFPQGKAREFSSEQTRANSPTRRELQVWGRDNNSLSEAGADRQGTVSFSFPQITLWQRPLVTIKIGGQLKEALLDTGADDTVLEEM
NLPGRWKPKMIGGIGGFIKVRQYDQILIEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEE
KIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLD
KDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELRQHLLRWG
FTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRGTKALTEVIPLTEEAELELAE
NREILKEPVHGVYYDPSKDLIAEIQKQGQGQWTYQIYQEPFKNLKTGKYARMRGAHTNDVKQLTEAVQKIATESIVIWGKTPKFKLPIQKET
WEAWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIVGAETFYVDGAANRETKLGKAGYVTDRGRQKVVSLTDTTNQKTELQAIHLALQDS
GLEVNIVTDSQYALGIIQAQPDKSESELVSQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRKVLFLDGIDKAQEEHEKYHSNWRA
MASDENLPPVVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKILLVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKTI
HTDNGSNFTSTTVKAACWWAGIKQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIVDIIA
TDIQTKELQKQITKIQNFRVYYRDSRDPLWKGPAKLLWKGEGAVVIQDNSDIKVVPRRKAKIIRDYGKQMAGDDCVASRQDED$
```

Fig. 110B

2003_CON_A2 pol.OPT

TTCTTCCGGCAGAGAACCTGGCCTTCCAGCAGCGGCGAGGCCCGCAAGTTCTTCCTCCGAGCAGCAGAGAACCGGCGGCCGCCAACTCCCCCACCTCCCGCGAGCTGCGCAACGG
CGGCCGCGACAACCTGCTGTCCGAGGCCGGCGCGCCGAGGAGCAGGCACCGTGCACTTCCTGCAACTTCCCCAGATCACCGTGGCCAGCGCGCCCCTGGTGA
CCGTGAAGATCGAGGGCCAGCTGCCGGAGGCCCTGCTGGACACCGGCGGCGCCCTGCTGGACATCAACCTGCCCGGCCAAGTGGAAGCCCAAG
ATGATCGGCGGCATCAAGTCGCGGCGGCTTCATCAAGGTGCCGCAACATGCTGGTGCAGCTGGGCTGCAGCCTGAACTTCCCCATCCCCCTGAACTTCCCCATCTCCCCGTGAACTTCCCCATCTCCCCGTGAAGC
CCGGCATGGACGGCCCCAAGGTGAACGGGCCCCTGACCGAGGAGCAGTGGCCCCTGACCGAGGAGAGATCAAGGCCCTGACCGAGATCTGCAAGGAGATGGAGAAGGAGGGCAAGATC
TCCAAGATCGGCCACCGAGAACCCCTACAACACCCCCGTGTTCGCCATCAAGAAGAAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAA
CAAGCGGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCCGCGAGGTGAAGAAGAAGAGTCCGTGACGTGGGCGACGCCT
ACTTCTCCGTGCCCCTGCACGAGGACTTCCGCAAGTACACCGCCTTCCTCCTCCAGATCGTTCCAGTGAGCCGCTCACCAGTACAACGTG
CTGCCCCCAGGGCTGGAAGGGGCTCCCCGGCCATCTTCCAGTCCATGACCAAGATCCTGGAGCCGGCCCTGCTGTCGCCGCTTCACCA
GTACATGGACGACCTGTACGTGGGCTCCGACCTGGAGATCGGCCAAGCGCACCGGCGCCAAGATGGACCTGCACCCGACAAGTGGCGCTGCCCGAGAAG
CCCCCGACAAGAAGCACCAGAAGAGCACCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCGACAAGCTGGGCATCAAGGTGAAGCTGTGCAAGCTGCT
GCGCGGCACCAAGGCCTGGAACGAGCCCTGACCGAGGAAGCTGGTTGACCCTGATCGCCGAGATCGGGCCAGCTGGAAGGGAGAACCTGACCCCGTGCACGGCG
TGTACTACGACCCCTCCAAGGACCTGATCGCCGAGATCGGCGCCGAGCTGGAGCTGGAGCTGAGCAGCCAGTGGACTGGACCCCTTCAAGAACCTGAAG
ACCGGCAAGTACGCCCGCATGCGCGGCGCCCACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATCGTGATCTGGGG
CAAGACCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGGGAGACCTGGGACCGGCGCCATCGCCGAGCTGGAGACCAAGCGCCGAGACCAAG
ACACCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCGCCCATCGCCGAGGACCTGGACCGGCGCCAGACCTTCTACGTGACGGGGCCGAGACCAAG
CTGGGCAAGGCCGGCTACGTGACCGACCGCGGCCGCCAGAAGATCGTGTCCCTGACCGAGACCACCAACCAGAAGACCGAGCTGCACGCCATCTACCTGGC
CCTGCAGGACTCCGGCCTGGAGGTGAACATCGTGACCGACTCCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCAGATCGGCGGCATCGCCGAGCAGCAGCTGGTG
TCCTCCCGGACATCGTGAAGCTGATCGAGAGGAGAAGCTGAAGAAGGAGAAGGTGTACCTGGCCTGGGTGCCCGCCCACAAGGGCATCGGCGGCAACGAGAGCTGGTG
CAACCTGCCCCCATCGTGGCCAAGGAGATCGTGGCCAGCTGCGACAAGTGCCAGCTGAAGGGCGAGGCCATGCACGGCCAGGTGGACTGCTCCCCCGGCA
TCTGGCAGCTGGACTGCACCCACCTGGAGGGCAAGGTGATCCTGGTGGCCGTGCACGTGGCCAGCGGCTACATCGAGGCCGAGGTGATCCCCGCCGAGACC
GGCCAGGAGACCGCCTACTTCATCCTGAAGCTGGCCGGCCGCTGGCCCGTGAAGGTGGTGCACACCGACAACGGCCCCAACTTCACCTCCGCCACCGTGAA
GGCCGCCTGCTGGTGGCCGTGCAGGAGTTCGGCATCCCCTACAACCCCCAGTCCCAGGGCGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGAGA
TCATCGGCCAGGTGCGCGACCAGGCCGAGCACCTGAAGACCGCCGTGTTCATCCACAACTTCAAGCGCAAGGGCGGCATCGGCGGCTAC
TCCGGCGGCGAGCGCATCATCGACATCATCGCCACCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCATCAAGATCCAGAACTTCCGCGTGTACTACCG
CGACTCCCGCGACCCCATCTGGAAGGGCCCCGCCAAGCTGCTGTGGAAGGGCGAGGGCGCCGTGGTGATCCAGGACAACTCCGACATCAAGGTGGTGCCCC
GCCGCAAGGCCAAGATCATCATCGCCGACTACGGGCGCGACGATGGCCAAGCAGATCATCAAGGCCAAGCTGCGCGTGGCCGCCGACGAGGAGGACTAA

Fig. 111B

2003_CON_B pol.OPT

```
TTGTTCCGCGAGGACCTGGCCTTCCCCCAGGGCAAGGCCCGCGAGTTCTCCTCCGAGCAGAGACCCCGCCCACCCGCCGCGGCCGAGCTGCAGGTGTG
GGGCCGCGACAACAACTCCCTGTCCGAGGCCCAGGCCACCGTCCTTCCTTCCCCAGATCACCCTGTGCGACGATCCCCGAGCGCCCCCCCTGGTGA
CCATCAAGATCGGCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCCACAAGGCCATCGGCACCGTGCTGGTGGGCCC
ATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAACCTGCTGACCCGAGATCGGCTGCCACCCCCCATCTCCCCATCGCCCCTGAACTTCCCCGTGAAGCTGAAGC
CACCCCCGTGAACATCATCGGCCGCAACATCCTGACCCAGATCGGCTGCCACCTCCCCCCATCTCCCCATCGCCCCTGAACTTCCCCGTGAAGCTGAAGC
CCGGCATGATGACGGCCCCAAGGTGAAGCAGTAGCAGTGGCCCCTGACCCAGGGAGAAGAGATCAAGGCCCTGGTGGAGATCTGCACCGAGATGGAGAAGGAGGGCAAGATC
TCCAAGATCGGCCCCGAGAATCGGCCACCCCTACACACCCCGTGTTCGCCATCAAGAAGAAGAGACTCCAAGTGGGCGAAGTCCGTGACCCGTGGGCGACGCCT
CAAGGCGCACCCCAGGACTTCTGGACAAGGACTTCCGGCAAGGCCTTCACCATCCCCTCACCTGACCCGGCATCCCCGCCATCCGCTACCAGACGTG
ACTTCTCCGTGCCCCTGGACAAGGACTTCCCGGCCATCTTCCAGTGCCTCCGCGGCCTGGAGCTGCACCACCTGCTGCGCGGCTTCACCA
CTGCCCCAGGGCTGGAAGGGTCTCCCCCGGCCATCTTCCAGTGCCTCCGCGGGCTGCTGCTGCAGCCCATCGTGCTGCCCCGAGAAG
GTACATGACGACCTGTACGTGGGCTCCGACCTGGAGATCGGCCAGCACAGCTGCGCCAGCCCATCGTGCTGCCCCGAGAAG
CCCCCCGACAAGAAGCACCAGAAGGAGCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGCCTATCAGGTGAAGCGCTGCT
GACTCCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCTCCAACTGGGCTGAGGAGGCCGTGCACGGCG
GCGGCGCAAGGCCCTGACCGAGGTGATCCCCCTGACCGAGGAGGCAGGGGCCCAGTGGACCGTGTGGGGCCCTTCAAGAACCTGAAGACCTGAAG
TGTACTACGACCCCTCCAAGGACCTGATCGCGGCCCGCCCCACCACCACCAGAAGGAGACCGCGTGGGGCGCGGCCCTTGGGG
ACCGGCCAAGTACGCCCGCATGCGGCGCGCCCATCAAGTTCAAGGCGTGGGAGGCCGTGAAGCAGTCGAGTCCCCGAGTCGGATCGTGGGG
CAAGACCCCCCAAGTTCAAGGCTGGGAGGCGCGCCATCAAGTTCAAGGCGTGGGAGGCCGTGAAGCAGTCGAGTCCCCGAGTTCGTGA
ACACCCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCGTGGTGGGCGCCCAACCGCGAGACCAAG
CTGGGGCAAGGGCTACGTGGGCTACGTGGGCTGTGGTACCAGCTGGAGGAGGAGGTCCAGCGCGGCCATCCACCTGGC
CCTGCAGGACTCCGGCCTGGAGGTGAACATCGTGACCGACTCCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACCAGAGTCCGAGCTGGTGT
CCCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGGCCTGGGTGCCCGCGCACAAGGGCATCGGCGGCAACGAGCTGGACAAGCTGGTG
TCCGCCGGCATCCGCAAGGTGCTGTTCCTGGACGGCATCGACAAGGCCCAGGACGAGCACGAGAAGTACCACCTGGGCGCCTCCGACTT
CAACCTGCCCCCGTGGTGGCTGACTGCACCCCACTTCCTGCTGAAGGGCGAGATCGGCGAGAGTTCGGCATCCCCTTACATGAGCAGCGCCTAC
GGCCAGGAGGAGACCGGCCAGGAGACCGCCTACTTCCTGCTGAAGCTGGCGGCCATCCACCACCGGCCAGGTGGAGTGATCCCCGCCGAGACC
GGCCAGGAGACCGCCTACTTCCTGCTGAAGCTGGCGCGCCATCCACCACCGGCCAGGTGGAGTGATCCCCGCCGAGACC
TCATCGGCCCGCGCGCACCAGCAGGCCGCATCCACCCGGCGCGGCACCTGGAGTGCATGAACAAGGAGCTGAAGAAGATCATCGGCCAGGTCCGCGACCAGCTGAAGACCGCTGCTGAAGAAGAGA
TCCGCCCGCGAGGCCCATGGTGGCCGCGGCATCGTGACCATCGGCGACATCGGCGACATCGCCGCCAACATCCAGAAGGAGATCAACAAGTGCGGCTAC
CGGACTCCCGCGCGTGGTGCGGTGGTGGAAGGGCCCGGGCTCGTGGAAGGGCCCTGCTGCCCGACCGAGCTGCAGCAGATGGCCGAGGGCCCAGGACATCCGCGACACAACTTCCGCGTGTACTACCG
GCCGCAAGGCCCAAGATCATCCGCGACTACGGCAAGCAGATGGCCGGCGACGACTGCGTGGCTCCCCCGCCAGGAGGCGAGGACTAA
```

Fig. 111C

67. 2003_B.anc pol.PEP

```
FFRENLAFPQGKAREFSSEQTRANSPTRRELQVWGRDNNPLSEAGADRQGTVSFSFPQITLWQRPLVTIKIGGQLKEALLDTGADDTVLEEM
NLPGKWKPKMIGGIGGFIKVRQYDQILIEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEE
KIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFERELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLD
KDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPEIVIYQYMDDLYVGSDLEIGQHRTKIELREHLLRWG
FTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRGTKALTEVVPLTEEAELELAE
NREILKEPVHGVYYDPSKDLIAEIQKQGQGQWTYQIYQEPFKNLKTGKYARMRGAHTNDVKQLTEAVQKIATESIVIWGKTPKFKLPIQKET
WEAWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIVGAETFYVDGAANRETKLGKAGYVTDRGRQKVVSLTDTTNQKTELQAIHLALQDS
GLEVNIVTDSQYALGIIQAQPDKSESELVSQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRKVLFLDGIDKAQEEHEKYHSNWRA
MASDFNLPPVVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFILKLLAGRWPVKVI
HTDNGSNFTSTTVKAACWWAGIKQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIVDIIA
TDIQTKELQKQITKIQNFRVYYRDSRDPLWKGPAKLLWKGEGAVVIQDNSDIKVVPRRKAKIIRDYGKQMAGDDCVASRQDEDS
```

Fig. 111D

2003_B.anc pol.OPT

```
TTCTTCCGCGAGAACCTGGCCTTCCCCAGGGCAAGGCCCGCGAGTTCTCCTCCGAGCAGCAGACCCGCGCCAACTCCCCACCCCGCGAGCTGCAGGTGTG
GGGCCGGCGGACAACAACCCCCTGTCCGGAGGCGGCGGCGGCCGGCCGGACCGGTGCTGCTTCCTTCCCCAGATCACCCTGTGGCAGGCACCGTGGCCCGCCAAGTGGAAGCCCAAG
CCATCAAGATCGGCGGCGGCCAGCTGGAAGGGAGGCCCTGCTGCGGCTTCATCAAGGTGCGCCAGATCCTGACCGAGATCGGCCACAAGGCCATCGGCCACCGTGCTGGTGGGCCCC
ATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAACCTGCTGACCCTGAACTTCCCCATCTCCCCATCGAGACCTGCACCGGAGAGGGCAAGATC
CACCCCGTGAACATCATCGGCGGCCAACTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGGTGGAGATCTGCACCGAGATGGAGAAGGAGGGCAAGATC
CCGGCATGGACGGCCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGGTGGAGATCTGCACCGAGATGGAGAAGGAGGGCAAGATC
TCCAAGATCGGCCACCCAGGACTTCTGGGAGGTGCAGCTGGCCATCCCCCACCCCGTGTTCGCCATCAAGAAGAAGGACTCCACCAAGTGGCGCAAGCTTCCGCGAGCTGAA
CAAGGCGCACCCAGGAGACTTCCGCGAGGGGCTCCCCGCCATCCCCCACCCCGTGTTCGCCATCAAGAAGAAGGACTCCACCAAGTGGCGCAAGCTTCCGCGAGCTGAA
ACTTCTCCGTGCCCCTGGAAGGGCTCCCCGCCATCCCCCCTTCCAGTGCTTCCAGTACCACCGCGCCTTCAGGAGCCCTTCCGCAAGCAGAGAAGCACCCCCAAGCAGAGAACCCCAAGCAGAGAACCCCAGTCACCGTG
CTGCCCCAGGGCTGGAAGGGCTCCCCCGCCATCTCCAGTGCTTCCAGTGCTTCCAGTACGACGCGTTCGCCATCAAGAAGATCCTGAGCCCTTCCATGACCAAGAGATCCTGAGCCCTTCCATGACCAAGATCCTGAGCCCTTCCATGACCAAGATCCTGATCTACCA
GTACATGACGACCTGTACGTGGGCTCCGACCTGGAGATCGGCCAGCACCGCAGCGCAGCTGCGGCGCGACCACCTGCTGCGCTGGGCTTCACCA
CCCCCGACAAGAAGCACCAGAAGGAGCACCCCCTTCCTGTGATGGCTACGACGTGCACCCCGACAAGTGCACCCGACGCGTTGCGGCGGCCCCAGAAG
GACTCCTGGACCGTGAACGAGCAGATCCAGAGAACGTGGCCTGAACGTGGGCGCAGCTGAACGGTGAAGGTGAAGCTGTGTCAAGCTGCT
GCGCGGCCAAGGCCCTGACCGAGGTGGTGCCCCTGACCGAGGAGGCCAGGGCCGTGGAGCTGACCGAGATCCTGAACGAGAGCCCGTGCACGGCG
TGTACTACGACCCCTCCAAGAGACCTGATCGCGCCGGCCGCCCAGGAGATCCAGAGAAGCAGGCCGGCCCAGGGCCCTTCAAGAACCTGAAG
ACGGGCAAGTACGCCCGCCATGCGCGGCGGCCACACCAACGACCTGAACGACCGTGAACGTGATCCGAGTCCCGCCGAGTCCCGAGTCCCGAGTGGGG
CAAGAGACCCCCCAAGTTCAAGCTGCCCATCCAGAGGAGGAGAACCAGAAGGAGGAGACCTGGGACGGCCGAGTACTGGCACGGCCACCTGACTGGCTGGTGA
ACACCCCCCCCAAGTTTCAAGCTGTGGTGTACCGTGGGGCGCCGCCAGAAGGTGGTGTCCCTGACCTACGTGGGGCGCGGCCATCGTGGGCGCCATCCACCTGGC
CTGGGCCAAGGGCCTGCGCTACGTGAAGGTGTGACCGACCGCGGCCGCCAGAAGGTGGTGTCCCTGACCTACGTGGGGCGCGGCCATCGTGGGCGCCATCCACCTGGC
CCCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACGGTGGGCCCCAGCTCCCAGTCCCGGTGCCCGCCAACGAGCAGGTGGAACAAGCTGGTG
TCCGCCGCTGCCCCCCGTGGTGACTGACGACGTGCTGTTCCTGGAGGGGCAAGATCATCCTGGCCCGCGTGCCCCGGTGGGCCCCGTGTCCATCGAGGGCCTCCCCGGCA
CAACCTGCCCCGTGGGCTGACTGACCGCCGCCAAGATCATCCTGGCCCGCGTGCCCCGGTGGGCCCCGTGTCCATCGAGGGCCTCCCCGGCA
TCTGGCAGCTGACCTGCAAGGTGCTGTTCTGCGACCTGGAGGGCAAGCTGCCCAAGGGCCTCCAACTTCACCTCCACCACCGTGAA
GGCCAGGAGACCGCCCTACTTCATCCTGGGGCCGCCCGTGAAGCTGGCCGCCAGGAGGTTGGGGAGTCCCATGAACAAGAAGCTGAAGAAGA
GGCCGCCTGCTGGTGGGCCAGCATCAAGGCCGAGTTCGGCGCCAGGCCACCTGCAGATGGCCGTGTTCATCCACCGTCAAGCGCAAGGGCGGCATCGGCGGCTAC
TCATCGGCGCGCGACCACCTGGCGACATCGTGCTGCACCGACATCATCGGCCAGAGCGTGCTGAGAAGCAGATCGCAGAAGATCCAGAAGACCAACTCCGCGTGTACTACCG
CGACTCCCGCGACCCCTGTGTGGAAGGGCCGGAAGGGCTGTGGGAAGGGCGGCGGCGACTGCGGTCAAGCTGCTGCTGTGTGATGGCCCGGCTACGACGAGGACGAGAGGACTAA
GCCGCAAGGCCCAGATCATCCGCGGCCTACGGCGACATCCGCGAGGGCCAAGCTGCTGCTGTGTGATGGCCCGGCTACGACGAGGACGAGAGGACTAA
```

Fig. 112A

68. 2003_CON_C pol.PEP

```
FFRENLAFPQGEAREFPSEQTRANSPTSRELQVRGDNPRSEAGAERQGTLNFPQITLWQRPLVSIKVGGQIKEALLDTGADDTVLEEINLPG
KWKPKMIGGIGGFIKVRQYDQILIEICGKKAIGTVLVGPTPVNIIGRNMLTQLGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEKIKA
LTAICEEMEKEGKITKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDEGFR
KYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRAQNPEIVIYQYMDDLYVGSDLEIGQHRAKIEELREHLLKWGFTTP
DKKHQKEPPFLWMGYELHPDKWTVQPIQLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGAKALTDIVPLTEEAELELAENREI
LKEPVHGVYYDPSKDLIAEIQKQGHDQWTYQIYQEPFKNLKTGKYAKMRTAHTNDVKQLTEAVQKIAMESIVIWGKTPKFRLPIQKETWETW
WTDYWQATWIPEWEFVNTPPLVKLWYQLEKEPIAGAETFYVDGAANRETKIGKAGYVTDRGRQKIVSLTETTNQKTELQAIQLALQDSGSEV
NIVTDSQYALGIIQAQPDKSESELVNQIIEQLIKKERVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGIDKAQEEHEKYHSNWRAMASE
FNLPPIVAKEIVASCDKCQLKGEAIHGQVDCSPGIWQLDCTHLEGKILVAVHVASGYIEAEVIPAETGQETAYYILKLAGRWPVKVIHTDN
GSNFTSAAVKAACWWAGIQQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIATDIQ
TKELQKQIIKIQNFRVYYRDSRDPIWKGPAKLLWKGEGAVVIQDNSDIKVVPRRKAKIIKDYGKQMAGADCVAGRQDED$
```

Fig. 112B

2003_CON_C_pol.OPT

```
TTCTTCCGCGAGAACCTGGCCTTCCCCCAGGGCGAGGCCCGCGAGTTCCCTCCGAGCAGCAGACCCGGCCAACTCCCCCACCTCCCGCGAGCTGCAGGTGCG
CGGCGACAACCCCGCTCCGAGGCCCGGCGGCCCGAGCGCCGCGCCGACGACACCCTGTGGCAGCACCCTGAACTTCCCCGAGCAGCCCCCTGTGTCCATCAAGGTGG
GCGGCCAGATCAAGGAGCCCTGCTGACACCGCGCGCCGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGCACCGTGCCCGTGGTGGGCCCACCCCGTGAA
ATCGGCGGCTTCATCAAGGTGCGCCAACATGCTGACCCAGTCTGCGAGATCCGCCCATCTCCCCATCCTGAACTTCCCGAGACCCGGCCGTGAAGCTGAAGCCCGGCATGGACG
CATCATCGGCGCCGCCAACATCTGACCCAGTCTGGCCTGACCGAGGAGATCAAGGCCCTGACCGGCCATCTGCGAGGAGATGGGAGAAGGAGGGCAAGATCACCAAGATCGGC
GCCCCAAGGTGAAGCAGTAGCCCCTGACCGGGAGATCAAGAAGAAGGACTCCACCAAGTGGCGCCAAGTGGTGACTTCCGCGAGCTGACAACAAGCGCACCCA
CCCGAGAACCCCTACAACACCCCGTGTTCGCCATCAAGAAGAAGGACTCCACCATCCCCTCCATCAACCGCCTTCACCATCCCCTCCATCAACCGCGCCCTACCTGCCCCCAGGGGC
GGACTTCTGGGAGGTGCAGCTGAACGCGCATCCCCGCCCAGTACACCGCCTTCACCATCCCCTCCATGACCAAGATCCTGAGCGTGCTGCTGCCCGCCTACTTCTCCGTGC
CCCTGGACGAGGGGCTTCCGCCAAGTACATCTTCCAGTACGCGAGATCGGCCAGCTGGCCTCCCAGATCTACCCCGAGAACCCCGACAAGCTGTGATCTACCAGTACATGGACGA
TGGAAGGGCTCCCCCGGCCATCTTCCAGTCCTGGAGATCGGCCAGCTGCGCGTGCGGGGCTTCACCACCCCGACAAGA
CCTGTACGTGGGCTCCGACCTGGAGATCGGCCACCGCCAGAGTGACCCCGAGCACCTGCGGCGCCCAGGGGCTTCACCACCCCGACAAGA
AGCACCAGAAGGAGCCCCTTCCTGTGATGGGCCAAGCTGACGAGCTGCACCCCGAGAATCCTACCCCGAGAATCCTACCCCGAGAATCCTACCAGAGAA
GTGAACGACACATCCAGAAGCTGGTGGGCCCCTGACCGTGCTGGAGCTGATCGGCCAGCTGCTGTGCAAGTGGCAAGTGGCAAGTGGCAAGTGGCAAGTGGCAAGTGGCAAGTGGCAAGACCCCCAA
GGCCCTGACCGACATCGTGCCCCTGACCGGCCAGCTGGACCGTGGACCCCCATCGCCGCGCCGAGAACCCGAGACTGCGAGAACCGGCGCCATCCGCCCTGCGAGGACTC
CCTCCAAGAGACCTGATCGCCCACCAGAGATCCAGAAGCAGCAGGCCACGAGGCCCTTCAAGAACCTGATCGAGACCGGCAAGTAC
GCCAAGATGCGACCGCCCACACCGCCGCCACGAGCGTGACGAGCCCGTGCAGAAGATCGACCGAGAAGATCGCCATGGCCAAGACCCCCCAA
GTTCCGCCTGACCGACGAGCTGGTGACCGTGACGCCGACTACTGGCAGGCCGGCATCGGCGAGTTCGTGAACACCCCCCCC
TGGTGAAGCTGTGTGGTACCAGCTGGAGAAGGAGCCCATCGCCGCGCCGAGAAGATCCGCGAGAGACCACCAGAGACTCGGCAAGGCC
GGCTACGTGACCGACGTGGCGCCGCCGCCAGAAGATCGTGTCCCTGACCGAGAAGATCTGCGAGCCCCTGGAGACTCGGCCCTGCGAGGACTC
CGGCTCCGAGGTGAACATCGTGACCGACGTGCTCCGTGTACCGACTCCCAGTACGCCCGGGTCCGCGGGTGCCCGAGGCCCAGTCGGCCCGTGAACCTGGTGTTCCTCCGGCATC
AGCAGCTGATCGAAGAGGAGCGCGTGTACCTGGACCGGCGCCGTGTCCTGGGGTGCCCCGAGGAGCAGGGCATCGGCGAGTTCGTGAGTTCAACCTGCCCCC
CGCAAGTGCTGTTCCTGGAACGCGCGCATCGACACGAGCGAGAGTACCACCTCCAACTGGCCCAGGTGGCCGCCATGGCCCAGGTGCTGCTCCCCCGGCATCTGGCCAGCTGG
CATCGTGTGGCACCCCGTGCGCAGATCGTGGCCAAGGCAAGATCATCCTGGTGGCCGTGCACGTGGCCGTCCCGGGATCTGGCCAGCTGG
ACTGCTGACCGAGGGTGAGGGCAAGCTGCTCGTGCGCGTACGTGACGTGTTCCTGTTCATCCACAACTTCAAGCGCAGGATCATCAAGGGCGGCATCCGCCAAGAGGACC
GCCTACTACTACGGCCGGACCGTGGGCGCGCCCATCCGTGTTCATCCTGTACGAGTTCCATGAACAAGGGCGGCGTGGTGAGTTCCATGAACAAGGGCGGCGTGGTGAGTTCCATGAACAAGATCGGCCGCAGG
GTGGCCGACCAGGCCCGGCACACCTGGAAGGTGCCCTCCGGCACCGGTGGTGATCCTCCACAACTTCCGCGCGGCATCGGCGGCTACTCCGCGACTCCCGCGA
CGCCATCATCGACATCATCGCCCACCGACATCCTCCAGGGCGGCCAAGATCGCCAGAAGAGCAGAACTTCAAGATCCGACAACATCCAGATCCTCCAGGACGAAACTCCCGACTCCCGCGA
AGATCATCAAGGACTACGGCAAGCAGATGGCCGGCGACGACTGCGTTGCTGTGAAGGTGCTGGGCCCCGCCGGGACTAA
```

Fig. 112C

69. 2003_C.anc_pol.PEP
FFRENLAFPQGEAREFPSEQTRANSPTSRELQVGRDNPRSEAGAERQGTLTLNFPQITLWQRPLVSIKVGGQIKEALLDTGADDTVLEEINL
PGKWKPKMIGGIGGFIKVRQYDQILIEICGKKAIGTVLVGPTPVNIIGRNMLTQLGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEKI
KALTAICEEMEKEGKITKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDEG
ERKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRAQNPEIVIYQYMDDLYVGSDLEIGQHRAKIEELREHLLKWGFT
TPDKKHQKEPPFLMMGYELHPDKWTVQPIQLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGAKALTDIVPLTEEAELELAENR
EILKEPVHGVYYDPSKDLIAEIQKQGHDQWTYQIYQEPFKNLKTGKYAKMRTAHTNDVKQLTEAVQKIAMESIVIWGKTPKFRLPIQKETWE
TWWTDYWQATWIPEWEFVNTPPLVKLWYQLEKEPIAGAETFYVDGAANRETKIGKAGYVTDRGRQKIVSLTETTNQKTELQAIQLALQDSGS
EVNIVTDSQYALGIIQAQPDKSESELVNQIIEQLIKKEKVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGIDKAQEHEKYHSNWRAMA
SEFNLPPIVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKVIHT
DNGSNFTSAAVKAACWWAGIQQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIATD
IQTKELQKQLIKIQNFRVYYRDSRDPIWKGPAKLLWKGEGAVVIQDNSDIKVVPRRKAKIIRDYGKQMAGADCVAGRQDED$

SUBSTITUTE SHEET (RULE 26)

Fig. 112D

2003_C.anc pol.OPT

TTCTTCCGCGAGAGAACCTGGCCTTCCCCCAGGGCGGAGGCCCGCGAGTTCCCTCCGAGCAGAGACCCGGCGCCAACTCCCCACCTCCCGGAGCTGCAGGTGGG
CCGCGACAACCCCGCTCCGAGGCGGCCGCCGAGGCGGCCAGGGCGGCCCGAGCGGCCCTGACCCTGAACTTCCCCCCTGTGGCAGCGCCCCCTGGTGTCCATCA
AGGTGGGCGGCCAGATCAAGGAGGAGGGCCCTGCTGGACGACACCGGCGGCCGCGCCAGTCCTGGAGGAGATCAACCTGCCCGGCAAGTGGAAGCCCAAGATGATC
GGCGGGCATCGGCGGCTTCATCAAGGTGCGCCGCCAGTACGACCAGAGTCCTGGGCGCAAGAGGCCATCGCCGGCAAGAGGGCCATCGAGACCGTGCCCGTGGGCGTGCCACCCCCA
CGTGAACATCATCGGCCGCAACATGCTGGCCCGTGAACATGCTGGCCCCTGACCCTGGGCTGCACCCAGCTGGCCCTGGGCGAGGAGGAGAAGATCAAGGGCCCTGACCGGAAGGAGGGCAAGATCACCAAG
TGGACGGGCCCCAGAGAACCCCTACACACACCCCGTGTTCGCCATCAAGAAGAAGAGACTCCACCAAGTGGCGAAGCTGGTGGGACTTCCGCGAGCTGAACAAGCG
ATCGGCCCCGAGAACCCCTACACACCCCGTGTTCGCCATCAAGAAGAAGAGACTCCACCAAGTGGCGCAAGCTGGTGGGACTTCCGCGAGCTGAACAAGCG
CACCCAGGACTTCTGGACGAGAGGGCTCCCCCGCCATCTCCAGCTGCAAGTGCAGCTGGCGTTCCGCCAAGTACACCCGCCTTCACCATGACCAAGATCCTGGAGTCCCGCGGCATCCCCCTCCACCGCCTTCACCATGACCAAGATCCTGGAGTCCCGCGAGATCGAGTGGGTTCACCACCCCCG
TCGTGCCCCTGACGAGAGGGCTCCCCCGCCATCTCCAGCTGCAAGTGCAGCTGGCGTTCCGCCAAGTACACCCGCCTTCACCATGACCAAGATCCTGGAGTCCCGCGAGATCGAGTGGGTTCACCACCCCCG
CAGGGCTGGAAGGGCCTCCCCCCCGGCTGCAAGTACACCGCCCTTCAGTCTGCGATGGGCTACGAGCTGAACTGGCCTCGGATGTGGGCAAGCTGGTGGGCCTGAACTGGCTGTGCCTGGTGCTGCGCGG
GGACGACCTGTACGGTGGTGGGCTCCCCCCCCCGGCTGCAAGTACACCGCCCTTCAGTCTGCGATGGGCTACGAGCTGAACTGGCCTCGGATGTGGGCAAGCTGGTGGGCCTGAACTGGCTGTGCCTGGTGCTGCGCGG
ACAAGAAGCACCAGAAGGAGCCCCCTTCCTGTGGGCAAGCTGGTGGGCCTGAACTGGCTGTGCCTGGTGCTGCGCGG
CGCCAAGGCCCTGACCGGACATCGTGTCCCTGACCGTGTCCCTGACCGTGTCCCTGACGTACCTGCTCGGGCAAGCTGGTGGGCCTGAACTGGCTGTGCCTGGTGCTGCGCGG
ACGACCCCTCCAAGGAGAGAGACCTGATCGCCGAGAAGAACGACGTGAAGCCCGCCCTTCAAGAACCTGAAGACCGGC
AAGTACGCCAAGATGCGCACCGCCCAACGACGTGAAGCCCGTGCAGAAGATCCGTGATCGTGGGGCAAGAC
CCCCAAGTTCCGCCTGCCCATCCAGAAGGAGAGACCTGGTGGACCGGCCACCTGGAGTTCGTGAACACCC
CCCCCCTGGTGAAGTGTGGTACCAGCTGGAGAAGGAGAGACCCATCCGCGGCCGCCAACCGCGAGACCAAGATCGGC
AAGGCCGGCTACGTGGACGCCGTGCTGGTGTGCCTGAGATCTGGGGCCGCGCCTGACCGAGAAGAGACCCAGCTGGCCCTGCA
GGACTCCGGCTCCGAGGTGAACATCGTGACCGACTCCCAGTACCGTCCCTGGGCGCCCACTGGCCATCCGAGCTGGTGAACCAGA
TCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGTCCCTGACAGACCGCGTGGGCGCCATGGGCCATGCCGAGTTCAACCT
GGCATCCGCAAGGTGCTGTTCCTGGACGCCATCCGCCTGGGCGCCATGCCGAGTTCAACCT
GCCCCCCATCGTGCTGGCCAAGGAGATCGTGGCCTCCTGCGACAAGATCATCGTGGGCCGCCAGTTGGACTGCGCCAGTTGGACTGCTGCCCCGGCATCTCTGGC
AGCTGGACTGCTGGCCACCCAGCTGGAGGCAAGATCATCGTGGGCCGCCAGTTGGACTGCCAGCTGCACCCTGGAAGGCCGCCAGTGGACTGCTGCCCCGGC
GAGACCGGCCTACTTCATCGTGGGCGCAATCTGGAGGGCAATCTCCAGCAGGCCGGTGGTGTCATCCACAACCCCAGTGGTCATCGCGCCGTGAAGGCCGC
CTGCTGGTGGTGCGGCGACCAGGCCGCGAGCCTTGGTGGGCATCGCCGCCTGGTGATCCCCGCCTACACCCGGCCGTGAAGAGATCATCG
GCCAGGTGCGCGACCAGGCCGATCATCGCCAAGGGCGCCGTCATCCACAACTTCAAGGGCCGGCATCCGAAGGCCGC
GGCGAGCGCATCATCGACACCGACCAGGAGAGCTGCAAGGCCCCAAGCTGCTGGAGGGCGGCCTGGTGCTGTGGAAGGCCCGCCGAGCTGCAAGGCCGC
CCGCGACCCCATCTCCGACTACGCAGCAGCAAGCGCCAAGCAGCAATCCGCGCCGACAACTCCGACATCAAGGTGGTGCCCCGCGCA
AGGCCCAAGATCATCCGCGACTACGGCAAGCAGAGATGGCCGCGGCCCCGACTGCCTGCCCGCCGACTGAGGACTAA

Fig. 113A

70. 2003_CON_D_pol.PEP

FFRENLAFPQGKAGELSSEQTRANSPTSRELRVWGGDNPLSETGAERQGTVSFNFPQITLWQRPLVTIKIGGQLKEALLDTGADDTVLEEIN
LPGKWKPKMIGGIGGFIKVRQYDQILIEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEK
IKALTEICTEMEKEGKISRIGPENPYNTPIFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDE
DFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPEIVIYQYMDDLYVGSDLEIGQHRTKIEELREHLLRWGF
TTPDKKHQKEPPFLWMGYELHPDKWTVQPIKLPEKESWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGTKALTEVIPLTEEAELELAEN
REILKEPVHGVYYDPSKDLIAEIQKQGQGQWTYQIYQEPFKNLKTGKYARMRGAHTNDVKQLTEAVQKIAESIVIWGKTPKFRLPIQKETW
ETWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIIGAETFYVDGAANRETKLGKAGYVTDRGRQKVVPLTDTTNQKTELQAINLALQDSG
LEVNIVTDSQYALGIIQAQPDKSESELVSQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSNGIRKVLFLDGIDKAQEEHEKYHNNWRAM
ASDFNLPPVVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKVILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKVVH
TDNGSNFTSAAVKAACWWAGIKQEFGIPYNPQSQGVVESMNKELKKIIGQVDSDIKVVPRRKVKIIRDYGKQMAGDDCVASRQDED$
DIQTKELQKQIKIQNFRVYYRDSRDPIWKGPAKLLWKGEGAVVIQDNSDIKVVPRRKVKIIRDYGKQMAGDDCVASRQDED$

Fig. 114A

71. 2003_CON_F1_pol.PEP

FFRENLAFQQGEARKFPSEQTRANSPASRELRVQRGDNPLSEAGAERRGTVPSLSFPQITLWQRPLVTIKIGGQLKEALLDTGADDTVLEDI
NLPGKWKPKMIGGIGGFIKVKQYDHILIEICGHKAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEE
KIKALTEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLD
KDFRKYTAFTIPSVNNETPGIRYQYNVLPQGWKGSPAIFQCSMTKILEPRTKNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELREHLLKWG
FTTPDKKHQKEPPFLWMGYELHPDKWTVQPIQLPDKDSWTVNDIQKLVGKLNWASQIYPGIKVKQLCKLLRGAKALTDIVPLTAEAELELAE
NREILKEPVHGVYYDPSKDLIAEIQKQGQGQWTYQIYQEPFKNLKTGKYAKMRSAHTNDVKQLTEAVQKIALESIVIWGKTPKFRLPIQKET
WDTWWTDYWQATWIPEWEFVNTPPLVKLWYQLETEPIVGAETFYVDGASNRETKKGKAGYVTDRGRQKVVSLTETTNQKAELQAIHLALQDS
GSEVNIVTDSQYALGIIQAQPDKSESELVNQIIEQLIQKEKVYLSWVPAHKGIGGNEQVDKLVSAGIRKILFLDGIDKAQEEHEKYHNNWRA
MASDFNLPPVVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKII
HTDNGSNFTSAAVKAACWWAGIQQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIA
TDIQTRELQKQITKIQNFRVYYRDSRDPVWKGPAKLLWKGEGAVVIQDNSEIKVVPRRKAKIIRDYGKQMAGDDCVAGRQDED$.

Fig. 113B

2003_CON_D pol:OPT

```
TTCTTCCGGCGAGAACCTGGCCTTCCCCCAGGGCCAAGGCCGGCGAGCTGTCCTCCGAGCAGACCGGCACCTCCCCCACCTCCCGCGAGCTGCGCGTGTG
GGGCGGCGACAACCCCCGTCCGAGACCGGCGCGGCGCCGCAGGCCCTGCTCGACACCGTCGCCGACGACACCGTTCCCCAGATCACCCCTGTGGTGACCA
TCAAGATCGGCGGCAGCTGGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCTGATCGAGATCGGCGGAGGATCAACCTGCCCCGCCAAGTGGAAGCCCAAGATG
ATCGGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCCGAGATCGGCTGCAACCTGCTGATCGAGATCTGCGGGCCATCGGACACCGTGTCCCGTGTGGGCCCAC
CCCCGTGAACATCATCGGCCGACAACCTGCTGCCCCAGATCGGCTGCAACCTGAACTTCCGCAGACCGTGCCCGTGCCCGTGAAGCTGAAGCCCG
GCATGGACCGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGAAGATCAAGGCCCTGACCGAGATCTGCACCGAGATGGAGAAGGAGGGCAAGATCTCC
CGCATCGGCCCCGAGAACCCTACAACACCCCATCTTCGCCATCAAGAAGAAGGACTCCACCAAGTGGCGAAGCTGGTGGTGGACTTCCGCGAGCTGAACAA
GCGCACCAGGACTTCTGGGAGGTGCAGGTGCAGCTGGAGGTCCCGCAGTACAACCGACTACACCCCGGCTCCTACCAGTCGTGATCTACCAGTA
TCTCCGTGCCCCTGACGAGGGCTCCCCGCCATCTCCAGTCCTCATGACCAAGATCCTGGAGCCCTTCCGCAAGCAGAACCCCAAGATCGTGATCTACCACCCC
CATGGACGACGTGTACGTGGGCTCCGACCTGGCTCGCCAGCCTGCCGGAGAGCTGCGCGAAGATCGGCACCCGCCGCTGCGCCTGGGCTTCACCACCC
CCGACAAGAAGCACCAGGAGGAGCCCCCTTCCTGTGATGGGCTACGAGAGCTGCACCCCGACATCTACCCCGGCATCAAGTGCCGCCTCTGTGCAAGCTGCTGCG
TCCTGGACCGTGAACGACATCCAGAGAACGACATTCCAGGATGTGGGCAAGCTGAACTGGGGCTCGAACTGGGTGGGCCTCGTGCTGCAAGCTGCTGCG
CGGCACCAGGCCCTGACCGAGGTGATCCCCCTGACCGAGGAGCCCGAGCTGGCCGAGCTGGCCGAGATCCTGAAGGAGCCCGTGCACGGCGTGT
ACTACGACCCCTCCAAGGACCTGATCGCCGAGATCCTGCAGCTGCAGGGCCAGTGCGAGACTACTACCAGAGACTCAAGGAGAACCTGAAGACC
GGCAAGTACGGCGCCCGCCATGCCGGCCCCACACAGACGTGAAGCAGCGTGACCGAGTCCATCGTGATCTGGGGCAA
GACCCCCAAGTTCCGCCTGCCCTCCAGGGCCATCCAGAAGGAGACCTGGTGGACAACCTGGGCCGCGCCAACCGGCCCATCAACCTGGCCCT
CCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCCGGCCGGCGCGGCCAGAAGGTGGTGCCCCTGACCGACCAGACCCAGAAGTCCCCT
GGCAAGGCCGTACGTGACGAGCCGGGCGGCAGATTGGGCATCGTGACCGACTGTGACCGAGCTCCAGTCCGAGCTGGTGTCCC
GCAGGACTCCGGCGGTGAGGTGAACATCGTGACCGACAGGAAGGTGTACCTGGCCTGGGTGCCCGCCCACAAGGGCATCGGGCAACGAGCAGCTGGTGTCC
AGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGGCCTGGGTGCCCGCCCACAAGGGCATCGGGCAACGAGCAGCTGGTGTCC
AACGGCATCCGCAAGGTGCTGTTCCTGGACGGCATCGACAAGGCCCAGGACGAGCACGAGAAGTACCACACAACTGGCCCGCGGCCATGGCGCCGCGGC
CCTGCCCCCGTGGTGGCCAAGGAGATCGTGGCCTCCTGCGACAAGTGCCAGCTGAAGGGCGAGGCTATGCACGGCCAGGTGGACTGCTCCCCCGGCATCT
GGCAGCTGGACTGCACCCACCTGGAGGGCAAGGTCATCCTGGTGGCCGTGCACGTGGCCAGCGGCTTCATCGAGGCCGAAGGTCCCAAGGAGGCTGAAGGC
CAGGAGGAGCTGGCTGTGGGCCGCCTGCTGGCCCTGGGGATCGACAAGGTTCTGCTGCACCCGGGTGGCCCCAGGGCCGTGGTGGAGTCCATGAACAA
CGCCTGCTGGGCCGCCTGCCGGCCATCATGCGCCTGGCCCCGATCCAGAAGGGCCGCAGAACTTCCGCGGTGTACTACCGCGA
TCGGCCGAGCGCCATCATCGACACCATCAAGCCCGAGCTGGAGATGGCCGCTGGTTCATCCACCACTTCAAGCGCAAGGGGCATCGGGCGGCTACTCC
GCCGGCCGAGCGCCATCATCGACACCATCAAGCCCGAGCTGGAGATCATCATCAGAGAACTCCAGGACAAGTCCGACATCCAGAACTCCGACATCCAGGACAACCTGCGCGTGTACTACCGCGA
CTCCCCGCGACCCATCGGAAGGGCCCCGACTACCGCCAAGCAGATGGCCCCGGCCAAGCATCGCGCTGCTGCTGTGGAAGGCGCGGCGCCCTCCCCGCC
GCAAGGTGAAGATCATCCGGCACTACGGCAAGCAGCAGATGGCCCCGGCCTGCTGCTGCGGCGACTGCAAGCATCCAGCACCTGCTGGCCTGGGCCGACATCCCGGCAGGACTAA
```

Fig. 114B

2003_CON_F1 pol.OPT

TTCTTCCGGGAGAACCTGGCCTTCCAGCAGGGCGAGGCCCGCAAGTTCCCCTCCGAGCAGAGACCCGCGCAGGGCCGAGCCGCGGCCGCACCCTGCCTTCCCCAGATCACCCTGTGGCAGCCCCCTGGTGA
GCGGCGGCGACAACCCCCTGTCCGGAGGCCGGCCGCGAGCCGCGCGGGCCCTGCTGGACACCGGCGCGCCGACAGCCATCAACCTGCCCGGCAAGTGGAAGCCCAAG
CCATCAAGATCGGCGGCGGCATCGGCGGCGGCTTCATCAAGGTGAAGCAGTACGACCACCATCCTGATCGAGATCGGCGGCCATCCCCATCTCCCCATCCCCCATCGGCCGTGCCGTGAAGCTGAAGC
ATGATCGGCGGCGGCATCGGCGGCCGCAACATGCTGACCGCGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGATCTGCACCGAGATCTGCGGCAAGTGGCGCAAGAGGGGCAAGATC
CACCCCCGTGAACATCAAGGTGAACATCAAGGTGCCGCAACATGCTGACCGCGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGATCTGCACCGAGATCTGGCAAGTGGCGCAAGAGGGGCAAGATC
CCGGCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGAAGATCAAGGCCCTGACCGAGATCTGCACCGAGATCTGGCGCAAGTGGGACTTCCGGCGACGCCT
TCCAAGATCGGCCCCGAGAACCCTACAACACCCCGTGTTCGCCATCAAGAGAAGAAGGACTCCACCAAGTGGCCAAGCTGGTGGTGACTTCCGGCGACGCCT
CAAGGCCACCCCAGGACTTCGGGAGGTGCAGCTGGCATCCCCACCCCGCGGCCTTCACCATCCCTCACCGACTACACCCCCGGCATCCGGCATCCAGTACAACGTG
ACTTCTCCGTGCCCCTGACCAAGGAGCTCCCCCGCCATCTTCCAGTGCTCCATGACCACCGCAGGATCCGGCACCGTGCTGCCGCACCGTGCTGCCCGACAAG
CTGCCCCAGGGCTGGAAGGGCTCCCCCGGCCTCCGACCTGGAGATGGGCTACGGCGGCAAGTGCGGCCTGCGCCGGCCGTGCTGCTGCCCGACAAG
GTACATGGACGACCTGTACGTGGGCCTGAACGACATCCAGAACGACTGGGCCTCCCAGGCCGAGTGGGCCTGAGCTGGCCGAGGCCGTGCACGGCG
CCCCGACAAGAAGCACCAGAAGGAGCTCCCCCAGGATCTGGAGCTGCACCCCGACAGCTGCCGAGATCTACCAGGATCTACCAGGAGAACCTGAAG
GACTCCTGGACGGTGAACGACATCCAGAGCTGGTGGGCCCTGACCGCGCCCTGACCGAGATCCAGGATCTACCAGGATCTACCAGGAGAACCTGAAG
GCGCGGCGGCCAAGGCCCTGACCGTGCCCTGACCATCGTGTCCGCCGAGCGGCCGGGGCAGCTGAAGCAGTGAAGATGCCCCTGGAGTCCATCGTCGTGGGG
TGTACTACGACCCCCTCCAAGGACCTGATCGGCCGAGATCCAGGATCCAGGGCCAGGGCCAGTGGACCTACCAGGAGCGTGGTGCAAGCTGCT
ACCGGCAAGTGAACGACCTCCGCCCAACGACGTGAGGCACGTGAAGGAGACCTGGGACACCTGGGACTACTGGTGCCGCCACTGGGGAGTTCGTGTGA
CAAGACCCCCAAGTTCCGCCTGCCCATCTGAAGGACGTGGGCCATCGCAGGCCACCTGGAGGTGCCGCCTCCAACCGCGAGACCAAG
ACACCCCCCCGTTGCCCATCGCCCGAGTGGGCCTTCTACGTTGACGGCCGAGATCGTGCCTGACCGGCCGGCCTCCAACCGCGAGACCAAG
AAGGGCAAGGCCGGCTACGTGACCGACCGCGGCCGGCCGAGAAGTGGTGTCCCTGACCGGACCCAGCTGCAGGCCATCCACCTGGC
CCTGCAGGACTCCGGGCTCCGAGGTGAACATCGTGACCGACTCCGTGACCCGGGCATCATCCAGGCCCAGCAAGTCCGACAGTCCGAGCTGGTGA
ACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGTCCTGGGTGCCCGCCCACAAGGGCATCGGCGGCAACCAACAGCTGGTCGACCAAGCTGGTG
TCCGCCGGCTGGACCGCGCCCTACTTCATCCTGAAGCTGGCCGGCCGCGACAAGGGCCATCATCCACACCGGCGTCCCGCTGCTCCCCCGGCA
CAACCTGCCCCGTGGTGGCTGGCCAAGGACATCGTGCCCTCCTGCGCACAAGTGCCGCCCGTGAAGATCATCCACACCGGCGTGGAGTGAACAAGCGGCGGCTAC
TCTGGCAGCTGGACTGCACCCACTCCGCCGGCAGCGAGTTCGGCATCGGCAGCAACTTCACCTGCGGCTGATGAACAAGGAGCTGAAGAAGA
GGCCAGGAGAGACCGGCCGCCTACTTCATCCTGAAGCTGGCCGACATCGTCGCCTGTTCATCCAGCAGATCCAGAAGGGCGGCATCGGCGGCTAC
TCATCGGCGGCGTGCGCGGCGCCAGCTTCGGGCGCAGGCCTGAAGACCCCGCGCCACCGACATCCAGAGATCCAGAACTTCCGCGTGTACTACCG
TCCGCCGGCCGGCCGACCATCATCGACATCGACTACCGGCCCCGCCGGCATGCGCGTGGTGGTGGAAGGGCGGCATCGGGCGGCTAC
CGACTCCCGCGCAAGGCCGACTACCGCGGCAAGCAGCAGATGCCGGCGCTGCTTCGGCGGCGTGTGAAGATCCGAGATCCAGGACCAACTCCGAGATCCAGAGGACGAGGACTAA
GCCGCGAAGGCCAAGATCATCCGCGGCTACGGCAAGCAGCAGATGCCGGCGCTACGACTGCGCGGCCAAGGTGGTGCCCCC

Fig. 115A

72. 2003_CON_F2 pol.PEP

```
FFRENLAFQQGEARKFSSEQTRANSPASRELRVRRGDNSLPEAGAERQGTGSSLDFPQITLWQRPLVTIKVGGQLREALLDTGADDTVLEDI
NLPGKWKPKMIGGIGGFIKVRQYDQIPIEICGQKAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEE
KIKALTEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDEWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLD
KEFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRAKNPEIVIYQYMDDLYVGSDLEIGQHRTKIEELREHLLRWG
FTTPDKKHQKEPPFLWMGYELHPDKWTVQAIQLPDKSSWTVNDIQKLVGKLNWASQIYPGIRVKHLCKLLRGAKALTDVVPLTAEAELELAE
NREILKEPVHGVYYDPSKDLIAEIQKQGHDQWTYQIYQEPHKNLKTGKYARRKSAHTNDVKQLTEVVQKIATEGIVIWGKVPKFRLPIQKET
WEIWWTEYWQATWIPEWEFVNTPPLVKLWYQLETEPIVGAETFYVDGAANRETKLGKAGYVTDRGRQKVVPLTETTNQKTELQAIHLALQDS
GSEVNIVTDSQYALGIIQAHPDKSESELVNQIIEQLIQKERVYLSWVPAHKGIGGNEQVDKLVSTGIRKVLFLDGIDKAQEEHEKYHSNWRA
MASDFNLPPVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKII
HTDNGSNFTSTVVKAACWWAGIQQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIA
TDIQTKELQKQITKIQNFRVYFRDSRDPVWKGPAKLLWKGEGAVVIQDNNEIKVVPRRKAKIIRDYGKQMAGDDCVAGRQDED$
```

Fig. 116A

73. 2003_CON_G pol.PEP

```
FFRENLAFQQGEAREFSSEQARANSPTRRELRVRRGDSPLPEAGAEGKGAISLSFPQITLWQRPLVTVKIGGQLIEALLDTGADDTVLEEIN
LPGKWKPKMIGGIGGFIKVRQYDQILIEISGKKAIGTVLVGPTPINIIGRNMLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEK
IKALTEICTEMEKEGKISKIGPENPYNTPIFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDE
NFRKYTAFTIPSTNNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRTKNPEIVIYQYMDDLYVGSDLEIGQHRAKIEELREHLLRWGF
TTPDKKHQKEPPFLWMGYELHPDKWTVQPIQLPDKESWTVNDIQKLVGKLNWASQIYPGIKVKQLCKLLRGAKALTDIVPLTAEAELELAEN
REILKEPVHGVYYDPSKELIAEVQKQGLDQWTYQIYQEPYKNLKTGKYAKRGSAHTNDVKQLTEVVQKIATESIVIWGKTPKFKLPIRKETW
EVWWTEYWQATWIPEWEFVNTPPLVKLWYRLETEPIPGAETYYVDGAANRETKLGKAGYVTDKGKQKIITLTETTNQKAELQAIHLALQDSG
SEVNIVTDSQYALGIIQAQPDRSESELVNQIIEQLIKKEKVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGIDKAQEEHERYHSNWRAM
ASDFNLPPIVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKVIH
TDNGSNFTSAAVKAACWWANITQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIAS
DIQTKELQKQITKIQNFRVYYRDSRDPIWKGPAKLLWKGEGAVVIQDNNEIKVVPRRKAKIIRDYGKQMAGDDCVAGRQDED$
```

Fig. 115B

2003_CON_F2 pol.OPT

Fig. 116B

2003_CON_G pol.OPT

```
TTCTTCCGCGAGAACCTGGCCTTCCAGCAGGGCGAGGCCCGCGAGTTCTCTCCGAGCAGGCCCGGCCCAACTCCCCACCCCGGCCGGAGCTGCGCGTGCG
CCGCGGCGACTCCCCCCCCTGCCCCGAGGCCGGCGCCGAGGGCAAGGGCGCCCGGCCATCTCCCCCAGATCACCCTGTGGCAGGCCCCTGGTGACCG
TGAAGATCGGCGGCGGCCAGCTGATCGAGGGCCCTGCTGGGACACCGTGCTGGGAGGAGATCCTGACCGTGGAGATCAACCGCAAGTGGAAGCCCAAGATG
ATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCTGATCGAGATCTCCCCATCTCCCCGTGAACTTCCCCGTGCCCGTGAAGCTGAAGCCCG
CCCCATCAACATCATCGGCCGCAACATGTCGACCGCAGATCGGCTGACCCCCTGACCGAGAGATCAAGGCCCTGACCGAGAGATCTGCACCGAGATCTGCACCGAGATCGCACCGAGATCTCC
GCATGGACGGGCCCAAGGTGAAGCAGTCAGTGGCCCCTGAGGAGAGATCAAGGCCCTGACCGAGAGAAGCTCCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAA
AAGATCGGCCCCCGAGAACCCCTACAACACCCCATCTTCGCCATCAAGAAGAAGGACTCCAAGAAGAGAAGTCCGTGACCGTGGGCGACGCCTACT
GCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCCGGCCTTCACCATCGCCAAGATCCTGGÄGCCCTTCCGCACCAAGAACCCCGAGATCGTGATCTACCAGTA
TCTCCGTGCCCCTGGAAGGGCTCCCCCGCCATCTTCCAGTCCTGGTGGCCGCAAGCTGAACGCGCCGAGCTGGAGCTGTGCAAGCAGCTGTGCTGCG
CATGGACGACGCCTGTACGTGGGCTTCCGACATCGGCCAGTACGACACCGTGCCAGATCGGCCAAGATCGCGCGAGCACCTGCTGGGGGCTTCACCACCC
CCGACAAGAAGCACCAGAAGAGGAGCCCCCTTCCTGGATGGGCTACGAGCTGCACCCCGACAAGGTGACCGTGCAGCCCCGACATCTACCCCGGCCTGAACAAGGTGAACA
TCCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCTGAACTGGGCCCGGCCTGAGAGCTGGAAGCAGCTGTGCAAGCTGCTGCG
CGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGCTGCCAGCAGGGCCTGGGCCTGGAAGCCCAGAGCTCCTGAAGGAGCCCTACAAGGACCTGAAGACC
ACTACGACCCCTCCAAGGAGCTGATCGCCGCCCACCACCAACGACGCGGCCTGCAGGAGGTGGTGCAGCAGCTGAAGCAGACCTGGGAGTTCGTGAACA
GGCAAGTACGGCGGCTCCGCGCCATCCGCCGCCGACCCCCAAGGAGACCTGGGAGGTGTGGTGGACCGAGACTGGCAGCGCCACCTGGGAGTTCGTGAACA
GACCCCCAAGTTCAAGCTGCCCATCCAGAAGGAGACCTGGGAGACCGTGGACCCGGCCGAGCCCATCCCGGCCCAACCGCCGAGACCAAGCTG
CCCCCCCCCAAGTTCAAGCTGCCCATCGTGGTGACCGGTACCCCGGCCGAGCCCATCATCCCAGTACGCCCTGGCTGCAGAGGCCGAGCCCCGAGCCCT
GGCAAGGCCGGCTACGTGACCGACCGCGGCCGCAAGAAGGTGCGCCTGGAAGACCCGGCCCGCCTGGGCCATCATCCCGAGCTGGTGAACC
GCAGGACTCCGGCTCCGAGGTGAACATCGTGACCGACAGCCCAGTACGCCCTGTCCTGGGTGCCCCGACAAGGCCATCGGCGGCAACGAGCAGGCTGGTGTCC
AGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGTCCTGGGTGCCCCGGCCCAGAGGCCGCAGCACGAGGCGCTACTCCAAGGGCCTCCAACTTCAA
TCCGGCATCGGCGCTGTGCTCGTGCCCCGCCAAGGAGATCGTGCCGCAGAGGAGCGACGAGCAGGGCCTCCTGCGCGCCCCCCCCGGCCGCCATCT
CCTGCCCCCATCGTGGCCAAGGAGATCCTGGAGGGCCAAGATCATCCTCTGGTGGGCCCGTGCTCCGGCTGGACTGCTCTCCCCGGCCGAGACCGGC
GGCAGCTGGACTGCACCCACCTGGAGGGCAAGATCATCCTGGTGGCCGTGCTGGTGCCCGCTACATCGAGGCGGCCTACACCAACCGCGACGCCGTGAAGGC
CAGGAGACCGCCTACTTCATCCTGAAGCTGGCCGGCCGAGGAGCTGGCCGTGCATCCTCGCACCCAGTAGGGCCGTGGCGTAGAGGAGCTGAAGAAGATCA
TCGGCCTGGTGGGGCCAACATCACCCAGGAGTTCGGCGATCCGAGAGAGGGCCACTGAGCTGCAGAAGCTGCCAGAAGCAGATCCAGAAGCTTCCACAACTCACCGGCGGCTACTCC
GCCGGCGAGCGCATCATCGACATCGCCGGCCCGCCCCGGCTGTGTGAGAAGATCCCAGCAGAGCCAGATCCAGAACACGACAACAACGAGAATCAAGGTGCCCCCCGCC
CTCCCCGCCGACCCCATCTGGAAGCTGACTACGGCCAAGCCAGAGATGGCCGGCGTGCTGTGGAAGGGCGACGCTGCGGCCAGGAGGACGAGGACTAA
GCAAGGCCAAGATCATCCGCGACTACGGCCAAGCCAGAGATGGCCGGCGTGCTGGGCCGGCCAAGCGACTAA
```

Fig. 117A

74. 2003_CON_H pol.PEP

FFRENLAFQQREARKFSPEQARANSPTSRELRVRRGDDPLSEAGAEGQGTSLSFPQITLWQRPLVTVKIEGQLREALLDTGADDTVLEEINL
PGKWKPKMIGGIGGFIKVRQYEQVAIEICGKKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEKI
KALTEICIEMEKEGKISKIGPENPYNTPIFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVSVLDVGDAYFSVPLDKD
FRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPEMIIYQYMDDLYVGSDLEIGQHRAKIEELRAHLLRWGFT
TPDKKHQKEPPFLWMGYELHPDKWTVQPVKLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVKQLCKLLRGAKALTDIVPLTKEAELELAENR
EILREPVHGVYYDPSKDLIAEIQKQGQPDQWTYQIYQEPFKNLKTGKYAKMRTAHTNDVKQLTEAVQKIATESIVIWGKIPKFRLPIQKETWE
TWTEHWQATWIPEWEFVNTPHLVKLWYQLETEPIAGAETYYVDGAANRETKIGKAGYVTDRGKQKVVSLTETTNQKTELQAIYLALQDSGL
EVNIVTDSQYALGIIQAQPDKSESELVNQIIEELIKKEKVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGIDKAQEEHERYHNNWRAMA
SDFNLPPIVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKVILVAVHVASGYIEAEVIPAETGQETAYFLKLAGRWPVKMIHT
DNGSNFTSAAVKAACWWADIQQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLRTAVQMAVFIHNEKRKGGIGGYSAGERIIDIIATD
IQTKELQKQISKIQKFRVYYRDSRDPIWKGPAKLLWKGEGAVVIQDNSEIKVVPRRKAKIIRDYGKQMAGDDCVAGRQDED$

Fig. 118A

75. 2003_CON_01_AE pol.PEP

FFRENLAFQQGKAGEFSSEQTRANSPTSRKLGDGGRDNLLTEAGAERQGTSSSFSFPQITLWQRPLVTVKIGGQLKEALLDTGADDTVLEDI
NLPGKWKPKMIGGIGGFIKVRQYDQILIEICGKKAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPISPIDTVPVTLKPGMDGPKVKQWPLTEE
KIKALTEICKEMEEEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLD
ESFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRIKNPEMVIYQYMDDLYVGSDLEIGQHRTKIEELRAHLLSWG
FTTPDKKHQKEPPFLWMGYELHPDRWTVQPIELPEKDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRGAKALTDIVPLTEEAELELAE
NREILKTPVHGVYYDPSKDLVAEVQKQGQDQWTYQIYQEPFKNLKTGKYARKRSAHTNDVRQLTEVVQKIATESIVIWGKTPKFRLPIQRET
WETWWMEYWQATWIPEWEFVNTPPLVKLWYQLEKDPIVGAETFYVDGAASRETKLGKAGYVTDRGRQKVVSLTETTNQKTELHAIHLALQDS
GSEVNIVTDSQYALGIIQAQPDRSESEVVNQIIEELIKKEKVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLEDGIDKAQEEHERYHSNWRT
MASDFNLPPIVAKEIVANCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKVILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKVI
HTDNGSNFTSAAVKAACWWANVRQEFGIPYNPQSQGVVESMNKELKKIIGQVREQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIA
TDIQTKELQKQITKIQNFRVYYRDSRDPIWKGPAKLLWKGEGAVVIQDNSDIKVVPRRKAKIIRDYGKQMAGDDCVAGRQDED$

Fig. 117B

2003_CON_H pol.OPT

```
TTCTTCCGGAGAACCTGGCCTTCCAGCAGCGGCGAGGCCCGCGGCCAACTCCCCACCTCCCGCGAGCTGCGCGTGCG
CCGGCGGCGACGACCCCCGTCCGGAGGCCGAGGGCCGAGGGCCAGGGCCGAGGGCCAGGGCCACCTCCACCCTGTGGCAGCGCCCCCTGGTGACCGTGA
AGATCGAGGGCGACGGCCAGCTGCGCGTGCGCGAGGGCCCTGCTGGACACCGGCGACACCGGTGCTGCTGGAGGAGATCAACCTGCCCGGCAAGTGGAAGCCCAAGATGATC
GGCGGCGCTTCATCAAGTGCGCCAGTACGAGCAGGTGGCCATCGAGGAAGGCCATCGGCACCGTGCCCGTGAAGCTGAAGCCCGGCA
CGTGAACATCATCGGCCGCAACATCCTGACCCAGATCGGCTGCACCCTGCCAACTCTCCCCATCGAACTTCCCCATCTCCCCATCGAGAAGCTGGACGCCCGGCA
TGGACGGGCCCCAAGGTGAAGCACGCAGTGGCCCCTGACCCGGAGGAGAGAGATCAAGGCCCTGACCGAGATCTGCATCGAGATGCAGAAGGAGGCCAAGATCTCCAAG
ATCGGCCCCCGAGAACCCTACACACACCCCATCTTCGCCATCAAGAAGAAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGCG
CACCCAGGACTTCTGGGAGGTGCAGCTGGCATCCCCAGCTGGCTCCTCCGAAGAGAAGGAAGTCCGTGTCCGTGCTGGGCGACGCCTACTTCT
CCGTGCCCCTGGACAAGGACTTCCGCCAAGTACACACGCGCCTTCACCATCCCGCCATCAACGAGACCCCCGGGGAGCTGTGGCCCGTGCAAGCGTGCTGCGCCC
CAGGGCTGGAAGGGCTCCCCCGCCATCTTCCAGTGCCTCCATGACCAAGATCCTGGAGCCTTCCGCAAGCAGAACCCCGAGATGATCATCTACCAGTACAT
GGACGACCTGTACGTGGGCTCCGACCTGGAGATCGGCCAGCACCGCACCAAGATCGAGGAGCTGCGCCAGCATCGAGGCCTTCACCACCCCG
ACAAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCGACATCTGGGCGTTGGGCAAGCTGAAGTGAAGCTGCAAGCAGCTGCTGCGCGG
TGGCCGGTGAACGACATCCAGAAGCTGGTGGGCCCCTGACCAGCTGGAGCTGGAAGCTGGGGAGTTCGTGAACACCC
CGCCAAGGCCCTGACCGAGGTCGTGCCCCTGACCGAGGAGGAGCAGGGCCCCATCGGCCGGGCCGAGCCTACTACGGCCGAGGCGCCCAACCGGCCATCTCCACCTGGCCCTGCA
ACGACCCCTCCAAGGACCTGATCGCCGGAGACCTGCAGGAGCTCCAGATCCAGAGAAGCAGGGCCCCTTCAAGAACCTGAAGACCGGC
AAGTACGCCAAGATGCGCCCTGCCCCCACACCGGCCCCATCCAGCTGAAGCAGGCCGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCGAGTCCATCGTGATCTGGGGCCAAGAT
CCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGGGAGACCTGGTGGACCGAGTATCCCGGATCCCGAGTGGGAGTTCGTGAACACCC
CCCACCTGGTGAAGCTGTGGTACCAGCTGGAGAAGAGCCCCGAGCGCGGCGCCGAGACCAACCGGCCATCTCCACCTGGCCCTGCA
AAGGCCGGCTACGTGACCGATGAAAGGTGGTCGTCCCTGACCGAGACCAGAGAAGGTGGTGTCCCTGACCGAGACCAGAGAAGTGGTGTCCAGTTCCGGGCCATCGTGGCCCTGCCCCGGAGAACGAGATCGGC
GGACTCCGGCAAGGCTACGTGACCGATGAAATCGTGACCGACTCCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCTGACCGCCCGAGGAGACCCCGGAGACCGGCCAG
TCATCGAGGAGCTGATCAAGAAGGAGAAGGTGTACCTGTCCTGGGTGCCCGCCCCACACAAGGGCATCGGCGGCAACGAGCAGGTGGACAAGCTGGTGTCCTCC
GGCATCCGCAAGGTGCTGTTCCTGGACGGCATCGACAAGGCCCAGGAGGAGCACGAGAAGTACCACAGCAACTGGCGCGCCATGGCTAGCAAGAGCTGGCGCGCCATGGCCTCCGACTTCAACCT
GCCCCCCATCGTGGCCAAGGAGATCGTGGCCTCCTGCGACAAGTGCCAGCTGAAGGGCGAGGCCATGCACGGCCAGGTGGACTGCTCCCCCGGCATCTGGC
AGCTGGACTGCACCCACCTTCATCCTGAAGCTGGAGGCCAGATCCAGCAGCATCCTGGTGGCCGTGCACGTGGCCAGCGGCTACATCGAGGCCGAGGTGATCCCCGCCGAGACCGGCCAG
GAGACCGCCTACTTCATCCTGAAGCTGGCCGGCCGCTGGCCCGTGAAGACCATCCACACCGACAACGGCAGCAACTTCACCTCCGCCACCGTGAAGGCCGC
CTGCTGGTGGGCGCCATCCAGCAGGAGTTCGGCATCCCCTACAACCCCCAGTCCCAGGGCGTGGTGGAGTCCATGAACAAGGAGCTGAAGAAGATCATCG
GCCAGGTGCGCGACCAGGCCGAGCACCTGCGCACCGCCGTGCAGATGGCCGTGTTCATCCACAACTTCAAGCGCAAGGGCGGCATCGGCGGCTACTCCGCC
GGCGAGCGCATCGTGGACATCATCGCCACCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCACCAAGATCCAGAACTTCCGCGTGTACTACCGCGACTC
CCGCGACCCCATCTGGAAGGGCCCCGCCAAGCTGCTGTGGAAGGGCGAGGGCGCCGTGGTGATCCAGGACAACTCCGACATCAAGGTGGTGCCCCGCCGCA
AGGCCAAGATCATCCGCGACTACGGCAAGCAGATGGCCGGCGACGACTGCGTGGCCGGCGCAGATGGGCCCGGCGACGACTGCGTGCCCGGCGACTAA
```

Fig. 118B

2003_CON_01_AE pol.OPT

TTCTTCCGCGAGAACCTGGCCTTCCAGCAGAGGGCAAGGCCGGCGGAGTTCTCCTCCGAGCCAGGGCAGCAGACCCGGCGCCAACTCCCCCCACCTCCCGAAGCTGGGCGACGG
CGGCCGCGGCGACAACCTGGCTGCTGAGAGCCGGCGGCCGAGCGCCAGGGCCACCTCCTTCCTTCCCCCGAGATCACCCTGTGGCCAGCGCCCCTGGTGA
CCGTGAAGATCGGCGGCCAGCTGAAGGAGGAGGCCGGGCTTCATCAAGGTGCGCGCCAGTACGACCAGAGTCCTGACACCCGGCGCCAAGTGGAAGCCCAAG
ATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCGCCAACATGCTGACCGGCAGCAGTGGCCCCTGACCGAGATCTGCAAGGAGGATGGAGGAGGGCAAGATC
CACCCCCGTGAACATCATCGGCGGCAACATCATCGGCCGCAACATCATGCTGACCGGCAGCAGTGGCCCCTGACCGAGATCTGCAAGGAGGATGGAGGAGGGCAAGATC
CCGGGCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGATCTGCAAGGAGGATGGAGGAGGGCAAGCTGGAA
TCCAAGATCGGCCCAGAGAACCCCTACAACACCCGTGTTCGCCAGTCTGCCATCCCCGGCAAGCTCACCAAGTGGCGCAAGTTCCGCGGAGCTGAA
CAAGGGCACCCAGGACTTCTGGGAGGTGCAGCTGGCCATCCCCGGCAAGCTCACCAAGTGGCGCAAGTTCCGCGGAGCTGGGCGACGCCCT
ACTTCTCCGTGCCCCTGGACGAGTCCTTCCGCAAGTACACCGCCTTCACCATCCCCTCCATCAACGAGACCCCGGCATCCGCTACCAGTACAACGTG
CTGCCCCAGGGCTGGAAGGGCTCCCCCGCCATCTTCCAGTCCTCCATGACCAAGATCCTGGAGCCCTTCCGCATCAAGAACCCCACCTGCTGTCCTGGGCTTCACCA
GTACATGGACGACCTGTACGTGGGCTCCGACATGGGCATCGGCCAGCTGGAGATCGGCCAGGCTGCACCCGGCTACGACGTTCACCCCCGAGAAG
CCCCCGACAAGAAGCACCAGAGGAGGAGCCCCTTCTGTGGATGGGCAAGCTGAACTGGGCCAGCTGGAGGGCCCCGTGCCCCGTGCACGGCG
GACTCCTGGACCGTGAACGACCCTGACCGAGCATCGTCGCCTGACCGGCCAAGCTGGAAGCTGGGGGCCCGAGTCTGATCTGGGG
GCGGCGGCAAGGCCCTGACCGAGGTCAGGGCCCAACGACGTGCGCGAGAAGATCGGCCACCTGGATCGCGATCCTGGGGGAGTTCGTGA
TGTACTACGACCCCTCCAAGGACCCTGGTGGCCGCCCACCAGCCGAGGTGCAGAAGATCGCCCACCTGGATCGCGATCCTGGGGGAGTTCGTGA
ACCGGCAAGTACGCCCAAGGCCCCGCCCCTGCCCCATCCAGCAGCGCGCTGGGTGGATGGAGTACTGGCAGGCCAGGGATCCCCGGCGCCTCCCCGAGACCAAG
CAAGAGACCCCCAAGTTCCGCGTGCGCCTGCCCCATCCAGCAGCGCGCTGGGTGGATGGAGTACTGGCAGGCCAGGGATCCCCGGCGCCTCCCCGAGACCAAG
ACACCCCCCCCCCTGGTGAACGTGCTGACCGACATCCAGAAGCTGGTGGCCATCTGGCTGGAGAAGGACCCCATCGTGCCCTGACCGAGCGCCTCCCCGAGACCAAG
CTGGGCAAGGCCGGCTACGTGACCGACCGCGGCAGGCAGGGCGCCATCTGGCTGGAGAAGGACCCCATCGTGCCCTGACCGAGCGCCTCCCCGAGACCAAG
CCTGCAGGACTCCGGCTGGAGCTGCCGAGGTGAACATCGTGACCGACGTGAACATGCTGACCCAGAAGGGGCATCGGCCACCTGCGCACCGACTT
TCCTCGGCATCCGCAAGGTGCTGTTCCTGGACGGCATCGACGGCATCGACGGCCAAGGTGATCCTGGCCGGCATCGGCCACCATGGCCTCCGACTT
ACCAGATCATCGAGGAGCTGATCAAGAAGGAGAAGGTGTACCTGTCCTGGGTGCCCGCCCACAAGGGGCATCGGCCACCTACCACTCCACCTGGTG
CAACCTGCCCCCCATCTGTGCAGGATCGTGGCCAAGTGATCCTGGCCGCAAGGTGATCCTGGCCGGCAAGGTGGACAAGCTGGTG
TCTTGGCAGCTGGACTGCACCCACCTCCTGCTGCAAGGTGATCCTGGCCGGCAAGGTGTGGTGGTCCATGACTTCATGACGCAAGGGCGGCATCGGCGGCTAC
GGCCAGGAGACCGCCTACTTCCTGCTGCAAGTGGCCATCCCCTACAACCCCAGTGCGTGTGGTGGAGTTCGCCATGACGCAAGGGCGGCATCGGCGGCTAC
TCATCGGCCAGGTGCTGGTGGCCGTGCCCGAGCCACCTGCCAGCTGCAAGGTGAATGCCGTGTTCATCCAGAAGGCGGCATCGGCGGCATCGGCGGCTACTACCG
TCCGCCCGGCCAAGGCCATCATCGACAATCCGCGCCAAGCTGCTGTTGATCCTGTGGTAGTCGAGAAGCCAGAACTTCCGCGTGTACTACCG
CGACTCCCGCGCCAAGGATCATCCGCGACTACGCCAGCAGATGCCGGCGGCGTGCTGTGGAAGGGCCGGCGACTCCGACATCAAGGTGGTGTGCCCC
GCCGCAAGGCCCAAGATCATCCGCGACTACGCCAGCAGATGCCGGCGGCGTGCTGTGGAAGGGCCGGCGACTGGCGGAGGACTAA

Fig. 119A

```
76. 2003_CON_02_AG_pol.PEP
FFRENLAFQQGEARKFSSEQTGTNSPTSRELWDGGRDNLLSEAGTEGQGTISSFNFPQITLWQRPLVTVRIGGQLIEALLDTGADDTVLEEI
NLPGKWKPKMIGGIGGFIKVRQYDQILIEICGKKAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEE
KIKALTDICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLD
KDFRKYTAFTIPSVNNETPGIRYQYNVLPQGWKGSPAIFQASMTKILEPFRTKNPEIVIYQYMDDLYVGSDLEIGQHRAKIEELREHLLRWG
FTTPDKKHQKEPPFLWMGYELHPDKWTVQPIQLPEKDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRGAKALTDIVTLTEEAELELAE
NREILKEPVHGVYYDPTKDLIAEIQKQGQGQWTYQIYQEPFKNLKTGKYAKMRSAHTNDVKQLTEVVQKVATESIVIWGKTPKFRLPIQRET
WEAWWMEYWQATWIPEWEFVNTPPLVKLWYQLEKDPIVGAETFYVDGAANRETKLGKAGYVTDRGRQKVVSLTETTNQKTELHAIHLALQDS
GSEVNIVTDSQYALGIIQAQPDRSESELVNQIIEKLIEKDKVYLSWVPAHKGIGGNEQVDKLVSNGIRKVLFLDGIDKAQEEHERYHSNWRA
MASDFNLPPIVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKVI
HTDNGSNFTSAAVKAACWWANVTQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIA
SDIQTKELQKQITKIQNFRVYYRDSRDPIWKGPAKLLWKGEGAVVIQDNSDIKVVPRRKAKIIRDYGKQMAGDDCVAGRQDED$
```

Fig. 120A

```
77. 2003_CON_03_AB_pol.PEP
FFRENLAFQQREARKFSSEQTRAISPTSRKLWDGGRDNPLPETGTEROQTASSFNFPQITLWQRPLVTVRIGGQLKEALLDTGADDTVLEDI
NLPGKWKPKMIGGIGGFIKVRQYDQILIEICGKKAIGTVLVGPTPVNIIGRNMLTQLGCTLNFPISPIETVPVTLKPGMDGPKVKQWPLTEE
KIKALTDICKEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLD
QDFRKYTAFTIPSTNNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPEIVIYQYMDDLYVGSDLEIGQHRTKIEELREHLLRWG
FTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYAGIKVRQLCKLLRGAKALTEVIPLTAEAELELAE
NREILKEPVHGVYYDPSKDLVAEIQKQGQGQWTYQIYQEPFKNLKTGKYARLRGAHTNDVKQLTEAVQKIATESIVIWGKTPKFKLPIQKET
WETWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIVGAETFYVDGAANRETKSGKAGYVTDRGRQKVVSLTDTTNQKTELQAIHLALQDS
GLEVNIVTDSQYALGIIQAQPDKSESELVSQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRKVLFLDGIDKAQEAHEKYHSNWRA
MASDFNLPPVVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFVLKLAGRWPVKII
HTDNGSNFISTAVKAACWWAGIKQEFGIPYNPQSQGVVESMNKQLKQIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIA
TDIQTKELQKQIIKIQNFRVYYRDSRDPIWKGPAKLLWKGEGAVVIQDNNDIKVVPRRKAKIIRDYGKQMAGDDCVASRQDED$
```

Fig. 119B

2003_CON_02_AG pol.OPT

```
TTCTTCCGGGAGAACCTGGCCTTCCAGCAGGGCGAGGCCCGCAAGTTCTCCTCCGAGCAGACCGGCACCAACTCCCCCACCTCCCCGCGAGCTGCTGGGACGG
CGGCCGCCGACAACCTGCTGCCGCCAGCTGATCGGCGGCCATCTCCTCCTTCAACTTCCCCAGATCACCTTGTGGGCCAGCCGGCCCCCTGGTGA
CCGTGCGCATCGGCGGCTTCATCGAGCTGCTGGACACCGGCGACCAGATCCTGATCGAGATCTGCGGCCAAGTGGAAGCCCAAG
ATGATCGGCGGCATCGGCGGCTTCATCGAGGTGCGCCAACATGCTGACCCCGTGAACATCATCGGCCGCCAACATGCTGACCCCCGTGAACTCATCATGCTGACCCCGTGAACTCGGCGGCAGATCTGCGGCCAAGAAGGCCATCGGCGAGACCGTGCCCGTGAAGCTGAAGC
CACCCCCGTGAACATCATCGGCCGCCAACATGCTGACCCCCGTGAACTTCCCCATCGAGCTGCGGCAGATCGGCCAACATGCTGACCCCGTGACCTGCCCGTGAAGCTGAAGC
CGGCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGACATCTGCACCGAGATGGAGAAGGAGGGCAAGATC
TCCAAGATCGGCCCAGAGAACCCCTACAACACCCCCGTGTTCGCCATCAAGAAGAAGGACTCCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAA
CAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCCGGCCTGAAGAAGAAAGTCCGTGACCGTGCTGGACGTGGGCGACGCCT
ACTTTCTCCGTGCCCCTGCCCGCACAAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCTCGATCAACCCCGGCATCCCGCTACCAGTACAACGTG
CTGCCCCAGGGCTGGAAGGGCTCCCCCGCCATCTTCCAGGCCTCCATGACCAAGATCCTGGAGCCCTTCCGCAAGAACAACCCCGAGATCGTGATCTACCA
GTACATGGACGACCTGTACGTGGGCTCCGACCTGGAGATCGGCCAGCACCGCAAGATCGAGGAGCTGCGCCAGCACCTGCTGCGCTGGGCCTTCACCA
CCCCCGACAAGAAGCACCAGAGAGAGCCCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCCAGCTGCCCGAGAAG
GACTCCTGACCAAGCGACGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCTCGAGGAGCCCGAGAAGCCCGTGCTGCCCGTGCACGGCG
GCGGCGGCCAAGGCCCTGACCGAAGTGATCCCGCTGACCGAGGAGGCCGAGCTGGAGCTGGCCGAGAACCGTGAGGAGCCCCTTCAAGAACCTGAAGC
TGTACTACGACCCCACCAAGAGACTTCGAGCTGCCCCAACACCAAGAGACGTGAAGCGGTGGTGCAGCAGTGGACCTACCAGGACGAGAAGGTGGCCAGGAGGCG
ACCGGCAAGTACGCGCAAGAATGCGCTCCGCCCACAACCAGAGATCCAGCGCGAGAAGACGTGAGCTGGTGCTGGGAGGTGGGGTTCGGTGA
CAAGAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGTCCACCTGGGAGGTGTGGTGGACCGAGTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTCGTGTGA
ACACCCCCCGTGAAGCTGTGGTACCAGCTGGGCGCCAACCGGCCCCATCGTGGGCGCCGAGACCTTCTACGTGGACGGCGCCGCCAACCGCGAGACCAAG
CTGGGCAAGGCCGGCTACGTGACCGACCGGGGCAGGCAGAAGGTGGTGTCCCTGACCGACACCACCAACCAGAAGACCGAGCTGCACGCCATCCACCTGGC
CCTGCAGGACTCCGGCTCCGAGGTGAACATCGTGACCGACTCCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCTGCCGGCAAGGCGGCATCAGCTGGTG
ACCAGATCATCGAGCAGCTGATCAAGAAGGACAAAGGTGTACCTGTCCTGGGTCCCGGCTCCGGCAGGAGCAAGGGCATCGGGGCGGCATCGAGCAGCTGGTG
TCCAACGGCATCCGCAAGGTGCTGTTCCTGGACGGCATCGACAAGGCCCAGGAGGAGCACGAGAAGTACCACTCCAACTGGCGCGCGCTCCGACTT
CAACCTGCCCCCGGTGGTGGCCAAGGAGATCGTGGCCTCCTGCGACAAGTGCCAGCTGAAGGGCGAGGCCATGCACGGCCAGGTGGACTGCTCCCCGGCA
TCTGGCAGCTGGACTGCACCCACCTGGAGGGCAAAGTGATCCTGGTGGCCGTGGCCGTGCACGTGGCCTCCGGCTACATCGAGGCCGAGGTGATCCCGGCCGAGACC
GGCCAGGAGACGCCTGCTGGTGGGCGCGCCACTTCATCCTGAAGCTGGCCGGCCGAGAGTTCGGCATCCCCTACAACCCCCAGTCCCAGGGCGTGGTGGAGAGCATGAACAAGGAGCTGAAGAAGA
GGCGCCCTGCTGGTGGGCGCGCCGACCAGGGCCAGGTCGGCATCATCCCCTACAACCCCCAGTCCCAGGGCGTGGTGGAGAGCATGAACAAGGAGCTGAAGAAGA
TCATCGGCCAGGTGCGCGACCAGGCCGAGCACCTGAAGACCGCCGTGCAGATGGCCGTGTTCATCCACAACTTCAAGCGCAAGGGCGGCATCGGCGGCTAC
TCCGCCGGCGAGCGCATCATCGACATCATCGCCACCGACATCCAGACCAAGGAGCTGCAGAAGCAGATCACCAAGATCCAGAACTTCCGCGTGTACTACCG
CGACTCCCGCGACCCCATCTGGAAGGGCCCCGCCAAGCTGCTGTGGAAGGGCGAGGGCGCCGTGGTGATCCAGGACAACTCCGACATCAAGGTGGTGCCCC
GCCGCAAGGCCAAGATCATCCGCGACTACGGCAAGCAGATGGCCGGCGACGACTGCGTGGCCAGCCGGCAGGACGAGGACTAA
```

Fig. 120B

2003_CON_03_AB pol.OPT
TTCTTCCGCGAGAACCTGGCCTTCCAGCAGCGGCGAGGCCCGCGAGGCCCGCGCCAGTTCTTCCTCCGAGCAGGACCCGGCGCCATCTCCCCCCACCTCCCGCCAAGCTGTGGGACGG
CGGCCGCGACAACCCCCTGCCCGACGCGGCCGGCACCGGCCGAGAGCGGCCTGCCCGACGACCGGCCTCCTTCAACTTCCCCAGATCACCTTCTGGGCAGCCGCCCCCTGGTGA
CCGTGCCGCATCGGCGGCGCCAGCTGAAGGAGGCCCTTCATCAAGGTGCGCCAGTACGACCCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCCCGTGGGCCCC
ATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAACATGCTGAGCCCAGCTGACCCAGCTGGGCTGGCCCTGAACTTCCCCATCTCCCCCACTCTGCAAGGCGCTGGCAAGC
CACCCCGTGAACATCATCGGCCGCCAACATGCTGAGCCCTGACCCAGCAGTGGCCCCTGACGGAGGAGATCAAGGCCCTGACCATCTGCAAGGCCGCAGATGGAGAAGGAGGGCAAGATC
CCGGCATGGACGGCCCCAAGTGAAGCAGTGGCCCCTGACCACCCCCGTGTTCGCCATCAAGGAAGAAGACTCCACCAAGTGGCCAAGCTGGTGGACTTCCGCGACGTGGGCGACGCCT
TCCAAGATCGGCACCCAGGACTTCTGGACCAGGGCTCCCCGCGCCATCCCCCTCCACCATCCCCGGCCTTCACCCGCCAAGTACACCGGCAACTCCCGCCAAGATCCTGAGCCCTTCACCA
CTGCCCAGGGCTGGAAGGGCTCCCCCGCCATCTTCCAGTCCTCGACCTGGAGATCGGCCAGCATCGAGGCCCAGCTGCCCGAGCACCTGCTCGCTGCCCCGAGAAG
GTACATGGACGACCTGTACGTGGGCTCCGACCTGGGCAAGCTGTGTGGGGACCAGCTGGTGCCCGAGATCTACGGCGAGAACTGGCCCAGCATCAAGGTGCGCCAGCTGTGCAAGCTGCT
CCCCCGACAAGAAGCACCAGAAGGAGCCCTTCCAGAGACTGGCCTCCCTGACCCCCGAGAGCTGGAACTGGGCTACGAGGCCCAGTTCCCCGAGATCTACCACCAGGAGATCGAGGAGTCCCTGACTCGGGG
GACTCCTGGACCGTGAACGACACATCCAGAAGGTGATCCCCCTGACCCCGGCGTGCTGGCCCCTGGCCGCGTGGGCCTTCAAGAACCTGATCTGGGG
GCGCGGCGCCAAGGCCCTGACCGAGGGTGATCCCCTGACCCCGCCGAGATCCCAGAAGCAGGCCCAGGCCTGGCGCAGCTGAACGACGTGGGCGCGCCCCACACCGGCCAAGAACCTGATCTGGGGA
TGTACTACGACCCCTCCAAGGGACCTGCCCGGCCGCGCCCCCCACCAGAAGCTGGAACGACACCTGGTGGCCCCCAGTTCGGTGTCCCCTGACCCCGCCGAGAGTCGTGCTGGGTGT
ACCGGCAAGTACCCCCAAGTTCAAGCTGCCCATCCAGAAGGAGACCTGGGAGACGACCTGGTGGCCTGGCGCTCCCAGTTCCGGCCCGACAAGCTGTGGTGT
CAAGACCCCAAGTTCAAGCTGCCCATCCAGAAGGAGACCTGGGAGACGGCGCCCAGAAGCTGCAGGCCATCACCGGCAAGCTGAGCGACTGGCAACGACAAGCTGGTG
ACACCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCGTGGGCGCGACCAAGACAGAAGAGACCCCGACAAGTGCCAACCGGCCGAGACCAAG
TCCGGCAAGGCCGGCTACGTGACCGACCGGGGCCGCCAGAAGTGGTGTCCCCTGACGCCGACCTCCCCAGTACGCCCCTGGGCATCATCCAGGGCCCCCGCCCGCGCGAGTCCGAGCTGGTGT
CCTGCAGGACCGCTCCGGCCTGGAGGTGAACATCGTGACCGAAGGAGAAGGTGTCCCCAGTACCTGGCCTGGCGACCTGCGGGCAACGCCAGCAGTGGACAAGCTGGTG
CCCAGATCATCGAGCGACGTGATCAAGGCTGATCAAGCCAGCCCGGCGGGCCATCTGGCCTCCGAAGATCATCCCCCCAGGCGCCATGCCTGAGTCGACTT
TCCGCCGGCATCGCAACCTGCCCCGTGTTCCTGACGCGCAAGATCTGTGGCCAAGGGCCAGTCGGCCATGCTGCTCCCCCCGGCA
CAACCTGCCCCGTGGTGGTGCACCCGCACTCTTCGTGCTGGCCCAGGGCCAAGATCATCCACATCTCCCGCTGTGGCCTGGCCTACTCGAGCTTCCACACCGCCGTGAACTCGAA
TCTGGCAGCTGGACTGGCACCCCTACTTCGTGCTGGCCCAGGGCCAGGGCATCGACTGAAGGAGGTTCGGCCCGTGAAGAACCCCCAGTCCCAGGTGGTGTCATCCAACGCAGCAGACAGA
GGCCGCCTGCTGGTGGGCACCGCCTACTTCGTGCTGGGGCCGCATCGTGCGCCGCCGATCAAGCCAGAGTTCGGCGCGTGTTCATCCACAACGCCAGCCGAGCTGAAGCAGGA
TCATCGGCCAGGTGCGCGACCATCATCGGCCCGCCATCATCGACACTGAAGCCCTGGAAGCGGCCGTGGTGTTCATCCACAACCACAAGCGGCGGGCTAC
TCCGCCGGCGACCTCCGCCGCCATCATCGACAAGGGCCAAGTGCTGCTGGTGGAAGGGGCCGACGACTGCGTGGCCCTCCCCCGCCGCCAGGGACTGA
GCCGCCAAGGCCAAGATCATCCGCGACTACGGCAAGCAGATGGCAGGCGCAGAAGGGCCGACTGCGTGGCCCGGCTCCCCGCCGCCAGGGACTGA

Fig. 121A

78. 2003_CON_04_CPX pol.PEP

FFRENVAFQQREARKFSSEQARANSPARRELRDERGDNLLSEAGTEGQGTISFNFPQITLWQRPLVTIKIGGQIREALLDTGADDTVLEEIN
LPGKWKPKMIGGIGGFIKVRQYDQIPIEICGKKAIGTVLVGPTPVNIIGRNMLTQLGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEK
IKALTEICTEMEKEGKISKIGPENPYNTPIFAIKKKNSTRWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDP
EFRKYTAFTIPSTNNETPGIRYQYNVLPQGWKGSPAIFQCSMTKILEPFRTKNPEIVIYQYMDDLYVGSDLEIGQHRAKIEELREHLLRWGF
STPDKKHQKEPPFLWMGYELHPDKWTVQPIQLAEKDSWTVNDIQKLVGKLNWASQIYPGIKVKQLCKLLRGAKALTDIVPLTTEAELELAEN
REILKEPVHGAYYDPSKDLIAEIQKQGQGQWTYQIYQEPYKNLKTGKYAKTRSAHTNDVRQLTEAVQKIAMECIVIWGKTPKFRLPIQKETW
DTWWTEYWQATWIPEWEFVNTPPLVKLWYQLETDPIAGAETFYVDGAASRETKQGKAGYVTDRGRQKVVSLSETTNQKTELQAIYLALQDSG
SEVNIVTDSQYALGIIQAQPDRSESDLVNQIIEQLIQKDKVYLSWVPAHKGIGGNEQVDKLVSNGIRKVLFLDGIDKAQEEHEKYHNNWRAM
ASDFNLPPVVAKEIVASCNKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKIIH
TDNGPNFTSAAVKAACWWADIQQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIAS
DIQTKELQKQITKIQNFRVYYRDSRDPIWKGPAKLLWKGEGAVVIQDNSDIKVVPRRKAKIIRDYGKQMAGDDCVAGRQDED$

Fig. 122A

79. 2003_CON_06_CPX pol.PEP

FFRENLAFQQGEAREFSSEQARANSPTRRELRVRRGDSPLPEAGAEGQGAISLSFPQITLWQRPLVTVRIGGQLIEALLDTGADDTVLEDIN
LPGKWKPKMIGGIGGFIKVRQYDQILIEICGKKAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEK
IKALTEICTEMEKEGKISKIGPENPYNTPIFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDE
DFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMIKILEPFRIKNPEIVIYQYMDDLYVGSDLEIGQHRAKIEELRELREHLLKWGF
TTPDKKHQKEPPFLWMGYELHPDKWTVQPIQLPDKDSWTVNDIQKLVGKLNWASQIYPGIKVKQLCKLLRGAKALTDIVPLTAEAELELAEN
REILKEPVHGVYYDPSKDLIAEIQKQGQGQWTYQIYQEPHKNLKTGKYARIKSAHTNDVKQLTEAVQKIALESIVIWGKTPKFKRLPIQKETW
ETWWTEYWQATWIPEWEFVNTPPLVKLWYQLETEPIVGAETFYVDGAANRETKKGKAGYVTDRGRQKVVSTGIRKVLVSTGIRKVFLDGIDKAQEDHERYHSNWRAM
SEVNIVTDSQYALGIIQAQPDKSESELVNQIIEQLIKKEKVYLSWVPAHKGIGGNEQVDKLVSTGIRKVFLDGIDKAQEDHERYHSNWRAM
ASDFNLPPIVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKVIH
TDNGSNFTSAAVKAACWWANITQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIAS
DIQTKELQKQITKIQNFRVYYRDSRDPIWKGPAKLLWKGEGAVVIQDNSEIKVVPRRKAKIIRDYGKQMAGDDCVAGRQDED$

Fig. 121B

2003_CON_04_CPX_pol.OPT

```
TTCTTCCGCGAGAACGTGGCCTTCCAGCAGCGCGGAGGCCCGCGAGGCCCGGCCCAACTCCCCGCCGCGCCGCGAGCTGCGCGACGA
GCGCGGCGACAACCTGCTGTCCGCGGAGGCCCGGCGCACCGAGGGCCAGGGCGACACCTTCCCCAGATCACCTGTGGCGCCCTGGTGCGCCCCTGGTGACCA
TCAAGATCGGCGGCGACCTGGGCGCCAGATCCGCGAGGGCCCTGCTGGACGACCGTGCTGCTGGAGGAGATCAACTTCGGGAGGAGGCCAAGTGGAAGCCCAAGATG
ATCGGCGGCGTTCATCAAGGTGCGCCAGATCCCCATCGAGGATCTCCCCATCTCCCCATCTGAACTTGCCCATGTGCCCGTGCCCGTGAAGCTGAAGCCCG
CCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGATCGGCTGCACCCTGAACTTCCACCGAGATCTGCACCGAGATGGAGAAGGAGGGCAAGATCTCC
GCATGGACGGCCCCAAGGTGAAGCAGTGCCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGAGATCTGCACCGAGATGGAGAAGTGGTGGCGAGCTGAACAA
AAGATCGGCACCCGAGAACCCCTACAACACCCCTACACCCCCACCCCGCCCATCCCCACCCCCGGGCATGGGCATCCGCCAGCTGGGGCGACGCCTACT
GCGCACCCAGGACTTCTGGGAGGTGCAGCTGGAGGTTCCGCAAGTACACCGCCTTCACCATGACCAAGATCCTGGAGCCCGAGATCGTGATCTACCAGTA
TCTCCGTGCCCCTGGACCCCGAGTTCCGCCATCTTCCAGTGCTCCATGACCAAGATCCTGGAGCCCGAGATCGTGATCTACCAGTAAGCTGGGGCCTTCTCCACCC
CATGGACGACGTGTACGTGGGCTCCGACCTGGAGATCGGCGAGCACCGTGCGCGAGCTGCGCGCCATCCAGCTGCGCAAGCTGCTGCG
CCGACAAGAAGCACCAGAAGGAGCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCGGACATCTACCCGGAGAACCGCGAGATCTTACCAGGAGCTGGGGCAA
TCCTGGACCGTGAACGAGCTGCCTGACCGAGAAGCTGGTGGGCCCTGACCGAGAGGCAGGGCGGCCAGTCCAGCGACGTGCCCAGGTGCATCGTGGGGCAA
CGGGCCCAAGGCCCTACGTGACCGAGATCGGCCCCACACACCGCTCCGGCCATCCAGAAGGAGACCCGGGACCTGGTGACCCGGGCCATCTACCTGGCCCT
ACTACGACCCCTCCAAGAGACCCGCTCCGCCCACACCAGAAGGAGACCGGGGGCCGTGACTGGCAGCCGGACCTGGTGACCCGGGCCATCTACCTGGCCCT
GGCAAGTACGCCCAAGAGAACCCGGTCCGCCATCCAGAAGGAGACCTGGGACCTGGTGACCCGGGCCATCTACGGCCATCTACCTGGCCCT
GACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACGACCTGGTGGACCGGCAGGGCCGCCCCTCCCCGCCATCTACCTGGCCCT
CCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCGTGGGCGTGTCCCCTGTGTCCGGACCGCGCATCGGCCCGCCGTGTGTCC
GGCAAGGCCGGCTACGTGACCGAGGTGAACATCGTGACCGCGCCGCCCATCCAGGGCCAGGGCATCGGCCCGCCGTGTGTCC
AGATCATCGAGACGCTGATCCAGAAGGAGATCGTGTTCCTGGACGCCCAGAGTCGTACCTGTCCGGACGCCCAGAGTCGGCTGGTGTCC
AACGGCATCCGCGGCTCCGAGGTGCTGTTCCTGGACGGCCAAGAGATCGTGACCGCGCCGCCCATGCGACTGCTCTCCCCGGCATCT
CCTGCCCCGTGGTGGCCACCCACCTGGCCTTCATCCTGAACGGCCTGGTGCCCGTGAAGCTGGGCCGGTGATCCCCCAGACCGGGC
GGCAGCTGGACTGCACCCACCTGGAGGGCCAAGAGTTCGGCCGCATCCCGACAACGGCGGCGTCCATGAACGGCATCAACCTCGCGTGAAGGC
CAGGAGGACGGCCTGGTGGCCGCGACATCCTGAAGGAGATCGTGCTGGGAAGATCCAGGACCAGAAGGGGCGCATCGGCGGCTACTCCGGA
CGCCTGCTGCCCCGACCAGCCGCCGACATCATCGACGACGGCCCGCCCCAAGGCCCAAGATCAAGCGGCAGAACTCCAGAACTCCGGA
TCGGGCCAGGTGTGGGCCGACGCCGCCAGGCCGCCGACATCATCGACCCAGCTGCGTGATCAAGGCGGCGGCGTGTACTACCGCGA
GCCGGGCGAGCGGCCATCATCGACACCGAGGGCCCCGGACCACTGCTGTGATCCAGGACACCAACTCCGACATCAAGGTGGTGCCCCGCC
CTCCCCGCCATCTGCGAAGGCCAAGATCAAGGGCGACTACGGCCAAGCTGCGTGGGCCGGCGACGAGATGGGCCAAGCAGCAGATGCGGAGGACTAA
```

Fig. 122B

2003_CON_06_CPX pol.OPT

```
TTCTTCCGCGAGAACCTGGCCTTCCAGCAGGGCGAGGCCCGCGCCAGTTCTCCTCCGAGTTCTCCTCCGAGCAGCAGCCCCGCCCCCACCCGCCGCCGCGCGAGCTGCGCGTGCG
CCGCGGCGACTCCCCCGCCCGAGGGCGGCGCCGAGGGCCAGGGCCAGAGGGCCCGCCCATCTCCCTGTCCTTCCCCAGATCACCCTGTGGCGCCCCCTGGTGACCG
TGGCCATCGGGCGGCCAGCTGATCGAGCTGCGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGTGGGCCCCCAC
ATCGGCGCATCGGGCGCGGCTTCATCAAGGTGCGCGGCTTCATCAAGGTGCGCGGCTTCATCAAGGTGCGCGGCTTCATCAAGGTGCGCCACCGTGCTGTGGGCCCCAC
CCCCGTGAACATCATCGGCCGCAACATGCTGCCTGCACCCTGACCCTGACCCTCCCCATCTCCCCATCTCCCCTGAACTTCCCCATCTCCCCATCTCCCCTGAAGCTGAAGCCCCG
GCATGGACGGGCCCAAGGTGAAGCAGTGGCCCCTGACCCTGACCGAGAGATCAAGGCCCTGACCGAGAGATCTGCACCGAGAGATGGAGAAGGAGGGCAAGATCTCC
AAGATCGGCACGGCCCCGAGAACCCCTACAACACCCCCATCTTCGCCATACCCTACAACACCCCCATCCACCCCCGCCTGAAGAAGAAGAAGACTGGTCCGCGAGCTGAACAA
GCGCACCCAGGACTTCTGGGAGGTGCAGCTGGCCATCCCCACCCCGCCTGCAGCTGCTGACCGTGGCCGACGCCTACT
TCTCCGTGCCCCTGGACGAGGACTTCCGCAAGTACACCCTTCACCATCCCTTCACCATGATCAAGATCCTGGAGTCGGAGGACTGCCACCACCC
CCCCAGGGCTGGAAGGGCTCCCCCGCCATCTTCCAGTCCTTCATGATCAAGATCTTCAAGATCCTGTGAGTGGGGCTTCCCGACAAGGAC
CATGGACGACCTGTACGTGGGCTCCGACCTGACCTGGGCTCCGACCTGACCTGACCTGACGGGGGGCCCCCTTCCTGTGGATGGGCAAGCTGAAGCTGCGCCGAGGGCCGAGCTGCGT
CCGACAAGAAGCACCAGAAGGAGGCCCCCTTCCTGTGGGCAAGCTGAAGCTGCGCCCTGACCGTGGGCTACCTCCGACACCACCAGATCTACCCCGAGAACGCGCCAGATCTACCCCGAGAACGCGGAGAATCTGATCTGGGGCAA
TCCTGGACGACCTGTACGTGGGCTCCGACCTGACCCTGACCCTGACGGGGGGCCCCCTTCCTGTGGGCAAGCTGAAGCTGCGCCCTGACCGTGGGCTACCTCCGACACCACCAGATCTACCCCGAGAACGCGCCAGATGGAGAAGGAGGGCAA
ACTACGACCCCCTCCAAGGACCTGATCGCCCCACCAACGACGTGAAGCAGTCCAGCCGGGGGCCCCACAAGAACCTGAAGACC
GGCAAGTACGCCCGGGCCCCCAAGTTCCCGGGCCCCGCCATCAGAAGGAGAGCTGGAGACCTGGGCTACTGGGGCTGGAGTTCGTGAACA
GACCCCCAAGTTCCCGGCTGGTGAAGTTCCCGGCTGGTGAAGTCTGGGCGCCATCAACCTGCCCCT
CCCCCCCCTGGTGAAGTCTGGGGCCCCGAGGTTCCTGACCGGCGCCTGGTGTCCCTGACCGAGAAGTGGTGTTCCCAGTACGCTCCCGAGCTCCGAGCTCCGACTTCAA
GGCAAGGGCGTACGGCGCCCGGACCCGGCCCCCCCCTGGTGTGTCCCTGACCTACCTGTTCCTGACCCGGCATCGAACATCGTGACCCCCCGGCATCGGCGCCATGGCGCCCTCCGACTTCAA
AGATCATCGAGCAGCTGATCAAGAGGAGAAGGTGTACCTGTCCCTGACGCGGCATCGGCGCCATGGGCGCGCCATGGCGCCAGTGGTGTCC
ACCGGCATCCGGCTCCGGCTCCGAGGTGATCGGCGCCATCGTGGCCGAGGAGACCACGAGCGCTACCACTCCAACTGGCCGCGCCTGAACTGCCCGAGCTGGTGAACC
CCTGCCCCATCGTGACCTGGCCAAGGAGATCGTGGCGCCAAGATCATCTGGCCTGGCCCTGACCTGGGCCCCGACAAGGCCCTCCCACCGACTGGGACTGCTCTCCCCCCGGCATCT
GGCAGCTGGACTGCACCCACTGGCCAAGGGCCAAGATCATCTGGCCTGGCCCTGGCGGCCTCCGGCCCTGACATGCGGCTACATCTGGGCGCTGTACCCCGCCGAGACCGGC
CAGGAGAGACCGGCCTGGCCCAAGATCATCTGGCCTGGCGGCCAAGGTGCGCCTGGCCCCTGAAGGTGATCCACCACCGGCCTCCAACTTCACCTGCGGCTACTCCCGGCGAAGGC
CGCTGCTGGTGGGCGCCCAAGATCATCCGGCCGACATCATCGCCTTCCGCGACATCATCGCCGGCTGGAGGTGATCCTGCGGCTGTTCATCAAGGCGGCTGGAGGTCAAGGCGGCTACTCCCGGCGA
TCGGCCCAGGTGCGCCGACCAGGCCGGGCCCCGGCCATCCCCTACAACCCGCCCGTGGCCCCGTGAAGGTGATCCCCTACAACCCGCCCGTGTTCATCCAGAAGCAGCTGCGCAAGGCGGCTACTCCCGGCGA
GCCGGCCAGGTGCGCCGACCAGGCCGGGCCCCGGCCATCGCCTTCCGCGACATCATCGCCGGCTGGAGGTGATCCTGCGGCTGTTCATCAAGGCGGCTGGAGGTCAAGGGCGGCTACTCCCGGCGA
CTCCCCGGCCGACCTGCTGTGGGAAGGGCCCCGGCCAAGCTGCTGGTGTGGAAGGGCCGACGACTTCCGCGTGGTGGTCCCGCC
GCAAGGCCAAGATCATCCGGCGGCTGCTGTGGGAAGGGCCGACGACTTCCGCGTGGGATCAAGGTGCTGGTGCCCCCCGCC
GCAAGGCCAAGATCATCCGGCGACTACGGCCAGATGCGCGGCGACTGCCTGGTGGGCGCCGAGGTGATCCTGCGGCTGTTCCTGGCCGCCATGGTGGTGCCCCCCGCC
GCAAGGCCAAGATCATCCGGCGACTACGGCCAGATGGGCGGCGACTGCCTGGTGTGGAAGGGCCGACGACTTCCGCGTGGTGGTCCCGCC
CTCGGCCCAGGTGCGCCGACATCATCGCCTTCCGCGACATCATCGCCCCGCCAAGCTGCTGGTGTGGAAGGGCCGACGACTTCCGCGTGGTGGTCCCGCC
GCAAGGCCAAGATCATCCGGCGACTACGGCCAGATGGGCGGCGACTGCCTGGTGTGGAAGGGCCGACGACTTCCGCGTGGTGGTCCCGCC
GCAAGGCCAAGATCATCCGGCGACTACGGCCAGATGGGCGGCGACTGCCTGGTGTGGAAGGGCCGACGACTTCCGCGTGGTGGTCCCGCC
CTCGGCCCAGGTGCGCCGACATCATCGCCTTCCGCGACATCATCGCCGGCTGGAGGTGATCCTGCGGCTGTTCATCAAGGGCGGCTACTCCCGGCGA
GCAAGGCCAAGATCATCCGGCGACTACGGCCAGATGGGCGGCGACTGCCTGGTGTGGAAGGGCCGACGACTTCCGCGTGGTGGTCCCGCC
GCAAGGCCAAGATCATCCGGCGACTACGGCCAGATGGGCGGCGACTGCCTGGTGTGGAAGGGCCGACGACTAA
```

Fig. 123A

80. 2003_CON_08_BC_pol.PEP

FFREILAFPQGEAREFPPEQTRANSPTSRELQVRGDNPSSEAGTERQGTLNFPQITLWQRPLVSIKVGGQIKEALLDTGADDTVLEEVNLPG
KWKPKMIGGIGGFIKVRQYEQIPIEICGKKAIGTVLVGPTPVNIIGRNMLTQLGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEKIKA
LTAICDEMEKEGKITKIGPDNPYNTPIFAIRKKDSSKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDKDFR
KYTAFTIPSVNNETPGIRYQYNVLPQGWKGSPAIFQCSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELREHLLKWGFTTP
DKKHQKEPPFLWMGYELHPDKWTVQPIQLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGAKALTDIVPLTEEAELELAENREI
LKEPVHGAYYDPSKELIAEIQKQGQDQWTYQIYQEPFKNLKTGKYAKMRTAHTNDVKQLTEAVQKIAMESIVIWGKIPKFRLPIQKETWETW
WTDYWQATWIPEWEFVNTPPLVKLWYQLEKDPIAGVETFYVDGAANRETKIGKAGYVTDRGRKKIVSLTDTTNQKTELQAIYIALQDSGSEV
NIVTDSQYALGIIQAQPDKSESELVNQIIEQLIKKERVYLSWVPAHKGIGGNEQVDKLVSNGIRKVLFLDGIDKAQEEHEKYHSNWRAMASD
FNLPPIVAKEIVASCDQCQLKGEAMHGQVDCSPGIWQLDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKVIHTDN
GSNFTSAAVKAACWWAGIQQEFGIPYNPQSQGVVESMNKELKKLIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIVDIIATDIQ
TRELQKQIIKIQNFRVYYRDSRDPIWKGPAKLLWKGEGAVVIQDNSDIKVVPRRKAKIIKDYGKQMAGADCVAGRQDEDS

Fig. 124A

81. 2003_CON_10_CD_pol.PEP

FFRENLAFQQRKARELPSEQTRANSPTSRELRVWGGDNTLSETGAERQGAVSLSFPQITLWQRPLVTVKIGGQLKEALLDTGADDTVLEEMN
LPGKWKPKMIGGIGGFIKVRQYDQILIEICGYKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEK
IKALTEICTEMEKEGKISRIGPENPYNTPIFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLYE
DFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPEMVIYQYMDDLYVGSDLEIGQHRIKIEELRGHLLKWGF
TTPDKKHQKEPPFLWMGYELHPDKWTVQPIQLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGAKALTDIVPLTEEAELELAEN
REILKEPVHGVYYDPSKDLIAEIQKQGQDQWTYQIYQEPHKNLKTGKYAKRRTAHTNDVKQLTEAVQKIAQESIVIWGKTPKFRLPIQKETW
ETWWTDYWQATWIPEWEFVNTPPLVKLWYQLEKEPIVGAETFYVDGAANRETKLGKAGYVTDRGRQKVISITDTTNQKTELQAINLALQDSG
SEVNIVTDSQYALGIIQAQPDKSESELVNQIIEQLIKKEKVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGIDKAQEEHEKYHNNWRAM
ASDFNLPPVVAKEIVASCDKCQLKGEALHGQVDCSPGIWQLDCTHLEGKVILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKVVH
TDNGSNFTSAAVKAACWWAGIKQEFGIPYNPQSQGVVESMNKELKKIIGQVVIQDNSDIKVVPRRKVKIIKDYGKQMAGADCVASRQDEDQ

Fig. 123B

2003_CON_08_BC pol.OPT

Fig. 124B

2003_CON_10_CD pol.OPT

TTCTTCCGCGAGAACCTGGCCTTCCAGCAGCGGCAAGGCCCGCGAGCTGCCCTCCGAGCAGAGACCCGCGCCCACCTCCCCGAGCTGCGCGTGTG
GGGCGGCGACAACACCCTGTCCGAGACGCGGCCGGCCGGAGCGCCCGGTGCCGGCCCGACGCACCGGCCCTGCTGGCCGCCAGTGGAAGCCAAGATG
TGAAGATCGGCGGCGCCCAGCTGGCTTCATCAAGGTGCGCCAGTACGACCCAGATCCTGATGATCGGCTACAACGGCCGTGCCCGTGAAGCTGAAGCCCG
ATCGGCCGTGAACATCATCGGCCCGACCATCGGCTGCTGACCCCTGACCCTGCACCTGCTGCACCCGAGATCGGCTGACCGAGATGGAGAAGGCAAGATCCC
GCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGAAGATCAAGGCCCTGACCGAGATCTGCACCGAGATCTGGTGGACTTCCGCGAGCTGAACAA
CGCATCGGCCACCCGAGAACCCCTACAACACCCCAGCTGGGCATCCCCACCCCGGCCTTCACCATCCCCATGACCAAGTCCGTGACCGTGGGCGACGCCTACT
GCGCACCCAGGACTTCTGGGAGGTGCAGCTTCCGCAAGTACAGACACGCGCCTCCTCCATGACCACCACCGCCTACGAGCTGCCCAGTACAACGTGCTG
TCTCCGTGCCCCTGTACGAGGGCTCCCCGGCTCCGACCTGGAGATCGGCCAGTACCTGGAGAGCTGCGGCCTTCACCACCC
CATGGACGACCTGTACGGCCGTCCGACCTGGAGATGGGCCAAGCTGAACCTGAACTGGCCTCCCCAGATCTACCCCGGCCATCAAGGTGCGCCAGTCTGCGCG
CCGACAAGAAGCACCAGAAGGAGCCCCCTTCTGTGGATGGGCCAAGCTGGGCAAGCTGAACTGGGCCTGAGCTGCCCGAGAAGGAC
TCCTGGACCGTGAACGACCCCTGAACGACATCCAGAAGGCCCTGACCGAGAGCGCCGAGCTGGAACTGGCCGAGCTGGGCCGTGCACGGCGTGT
CGGCGCCAAGGCCCTGACCGAAGCTGGCCGTGCCCCTGACCGAGATCCAGAGCGGCCTGGGCGAGATCTACCCCGGCGAGAACCGCGAGAACCTGAAGACC
ACTACGACCGTGAAGCACCGTGAACGACATCCAGAGGCGCCGCGCGACGAGGCTGGGCGAGATCGGCCCCCACAAGAACCTGAAGACCCTGATCGTGGGCAA
GGCAAGTACGACCCTGCCCTGACTGCCGCCAACGACATCCAGAAGGAGAGCGTGGGCTGCTGGGCTACCAGCTGGGCCAACCGCGAGACCAAGCTG
GACCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCGTGGGCGTGGGCCGCGCCGACACCAACCAGAGCTGCAGCCGCCATCAACCTGCCCT
GGCAAGGCCGGCTACGTGACCGACCGCGGCCGCCAGAAGGTGATCTCCATCACGCCCTGGGCCCCGCGACCTGCGAGCTGGTGAACC
GCAGGACTCCGGCCTCCGGAGGTGAACATCGTGACCGACCTGGTACCTGTCCTGACCATGCCCCGCCCGGTGCCCGCCAGGAGGAGCACGCGAAGGGCATGTCC
AGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGTCCTGGGTGCCCGCCCACAAGGGCATCGGCGGCAACGAGGTGGACAAGGCCTGTGTCC
TCCGGCATCGGCCCGTGGTGGCCAAGGAGATCGTGGCCTCCTGCGACAAGTGTCGGGCCAAGGTGGTCCTGCGGACGGCCGCCATCTGGCCCGGCATCT
CCTGCCCCCCGTGGTGGCACCCCCACCTGGCCCAAGTCGTGGCCGTGAAGCTGGCCGAGCTGGGCCTGGTGCTGGAGGGGCCGAGACCGGC
GGCAGCTGGACTGCACCCATCTGGAAGCTGGCGGCCCGGTGAAGTGCCCGGCCGTGAAGTGGTGCCGGGTGTGCTCCAGTCCCCCACAACCCCGAGTGGAAGGC
CAGGAGACCGCCTACTTCCTGCTGAAGCTGGCCGGCCGTGGGCCGGACATCCCTACACCAACCAGCAGGTGCTGGTGTCCAACTTCACCTCCGCCGTGAAGGC
CGCCTGCTGGCGGCCGGCTACGTGGGCCGGCCGAGCATCCCGCCACTACCAGGAGGTTCGGCATCAAGCAGGTGCCGGTGTTCATCCCACAACTTCAAGCGGCGGCTACTCC
TCGGCCCAGGTGCCCGGCCAGGCCGAGCACACCTGCACCCTGCCACCACCGCCGAGCTGCAGAAGGAGATCCAGAAGCAGATCATCAAGGTGCAGAACTTCCGCGTGTACTACCGCGA
GCCGGGCGAGCCGCCATCATCGACACCGCCCCAAGCTGCTGTGGAAGGGCGAGGGCGCCGTGGTGCTCCAGGACGAGGACCAG
GCAAGGTGAAGATCATCAAGGACTACGGCCAAGCAGCAGTACGGCGACGAGGACCAG

Fig. 125A

82. 2003_CON_11_CPX pol.PEP

FFRENLAFQQGEAREFSPEQARANSPTSRELRVRGGDSPLPETGAEGEGAISFNFPQITLWQRPLVTIKVAGQLKEALLDTGADDTVLEEID
LPGRWKPKMIGGIGGFIKVRQYEEIIIEIEGKKAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPISPIDTVPVKLKPGMDGPKVKQWPLTEEK
IKALTEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDE
SFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRTQNPEIVIYQYMDDLYVGSDLEIGQHREKVEELRKHLLKWGF
TTPDKKHQKEPPFLWMGYELHPDKWTVQPIQLPDKECWTVNDIQKLVGKLNWASQIYPGIKVKQLCKLLRGTKALTDIVPLTAEAELELAEN
REILKEPVHGVYYDPSKDLIAEVQKQGLDQWTYQIYQEPFKNLKTGKYAKRRTAHTNDVRQLAEVVQKISMESIVIWGKIPKFRLPIQRETW
ETWWTDYWQATWIPEWEFVNTPPLVKLMYQLEKEPIIGAETFYVDGAANRETKLGKAGYVTDKGRQKVVTLTETTNQKTELEAIHLALQDSG
LEVNIVTDSQYALGIIQAQPDKSESELVSQIIEQLIKKEKVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGIDKAQEEHERYHSNWRAM
ASDFNLPPIVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKVIH
TDNGSNFTSAAVKAACWWANIQQEFGIPYNPQSQGVVESMNKELKKIIGQVREQAEHLKTAVQMAVFIHNFKRRGGIGGYSAGERIVDIIAT
DLQTKELQKQITKIQNFRVYYRDSRDPIWKGPAKLLWKGEGAVVIQDNSDIKVVPRRKAKIIRDYGKQMAGDDCVAGRQDED$

Fig. 126A

83. 2003_CON_12_BF pol.PEP

FFRENLAFQQGEARKFPSEQARANSPASRELWVRRGDNPLSEAGAERRGTVPSLSFPQITLWQRPLVTIKVGGQLKEALLDTGADDTVLEDI
NLPGKWKPKMIGGIGGFIKVKQYDNILIEICGHKAIGTVLVGPTPVNIIGRNLLTQLGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEE
KIKALTEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLD
KDFRKYTAFTIPSVNNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELRQHLLRWG
FTTPDKKHQKEPPELWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVKQLCRLLRGTKALTEVIPLTKEAELELAE
NREILKEPVHGVYYDPSKDLIAEIQKQGQGQWTYQIYQEPFKNLKTGKYARMRGAHTNDVKQLTEAVQKITTESIVIWGKTPKFRLPILKET
WDTWWTEYWQATWIPEWEFVNTPPLVKLWYQLETEPIAGAETFYVDGASNRETKKGKAGYVTDRGRQKAVSLTETTNQKAELHAIQLALQDS
GSEVNIVTDSQYALGIIQAQPDKSESELVNQIIEQLIKKEKVYLSWVPAHKGIGGNEQVDKLVSAGIRKILFLDGIDKAQEEHEKYHNNWRA
MASDFNLPPVVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKIILVAVHVASGYLEAEVIPAETGQETAYFILKLAGRWPVKTI
HTDNGPNFSSAAVKAACWWAGIQQEFGIPYNPQSQGVVESMNKELKKIIRQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIS
TDIQTRELQKQIIKIQNFRVYYRDSRDPVWKGPAKLLWKGEGAVVIQDNSEIKVVPRRKAKIIRDYGKQMAGDDCVAGRQDED$

Fig. 125B

2003_CON_11_CPX pol.OPT

TTCTTCCGCGAGAACCTGGCCTTCCAGCAGGGCGAGGCCCGCGAGTTCTTCCCCGAGCAGCCCGCGCCAACTCCCCACCTCCCGCGAGCTGCGCGTGCG
CGGCGGCGACTCCCCCCTGCCCGAGACCGGCGAGGGCGAGGGCGGCGCCATCTCCTTCAACTTCCTCCAGATCACCCTGTGGCAGCGCCCCCTGGTGACCA
TCAAGGTGGCCGGCGCCAGCTGAAGGAGGCCCTGCTGGACACCGGCCTGAGAGATCGACGTCTGGAGGAGATCATCGAGATCAGGGCCAAGAAGGCCATCGACACCGTGCCCGTGGCGCTGAAGCTGAAGCCCG
ATCGGGGCCATCGGGCGCTTCATCAAGGTGCGCCAGCTACGAAGGTGCGCCGCAACATGCTGACCGCAACATCGGCCTGAACATGCTGACCCCTGAACTTCCCCGACTTCCCCATCTCCCCCATGCTGACCCCTGAACTTCCCCGAC
CCCCGTGAACATCATCGGCCGCAACATCGGCCCTGACCCCGAGGAGAAGTCAAGGCCCTGACCGAGATCTGCACCGAGATGGGAAGAGGGCAAGATCTCC
GCATGGACGGCCCCAAGGTGAAGCAGTCGCCCCTGACCGAGGAGAAGGAGGTGACTTCCGCGAGCTGAACAA
AAGATCGGCACCCCGAGAACCCCTACAACACCCCGTGTTCGCCATCAAGAGAAGGACTCCACCAAGTGGCAAGCTGGTGGACTTCGGCGACGCCTACT
GCGCACCCAGGACTTCTGGACGAGTCCTTCCGCAAGTACACCGCGCCTTCCATGACCAGATCCTGCCCGAGAACCCCGCACCCAGTACCAGTA
CCCCAGGCGTGGAAGGGCTCCCCCGCCCATCTTCCAGTCCTGCCAGCCGCTGGGCTCCGACATCGGCCAGAGCACCTGCTGAAGTGGGCTTCACCACCC
CATGGACAGACCTGTACGTGGGCTCCGAGATGGGCTACGAGCTGGATGGGCAAGCTGAACCCCGAGATCTACCCCGGCATCAAGTGAAGCAGCTGTGCGCG
CCGACAAGAAGCACCAGAAGGAGCCCCCTTCCTGTGGATGGGCTACGAGCTGTGGTGGGCAAGCTGAACTGGCCTCCGAGGACATCCCCGACAAGGAG
TGCTGGACCGTGAACGACGTCCCCTGACCGCTGCCCCGAGGCTGGTGGGCCCCTGACAAGTGAAGCAGTCGTGCAAGCTGCTGCG
CGGCACCAAGGCCCTGACCGACGACATCTGACCGACGCCGAGAGGTGCAGAGGTGCGGAAGGAGCCCGTGCCCGGCGTGTCGAGACC
ACTACGACCCCTCCAAGGACCCTGATCGCCCACCAACGACGTGCCGCCAGCTGGGTGCAAGAGATCTCCATCGTCGATCTGGGGCAA
GGCAAGTACGCCAAGCGCCGCCAAGCGCCGAGAACGCGCCACCTGGTGACCGCACTGGCGAGACCTGGATCCCGAGTGGGAGTTCGTGAACA
GATCCCCAAGTTCCGCCTGCCCATCCAGCGCGAGATCATCGGGCCCGCGCCATCATCGCCGAGAACCGCGAGACCAAGCTG
CCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGTGTACCGACAAGGGCCGCCAGAAGGTGGTGTACCAGCTGCTGGAGGCCCATCCACCTGGCCCT
GGCAAGGCCGGCTACGTGACCGAGAAGGGCCGCCCAGAAGGTGGTGACCGACTCCCAGTAGCCCCTGGGCTCCGCGAGCTGGTGGTGTCCC
GCAGGACTCCGGCTGGAGGTGAACATCGTGACCGACTCCCAGTAGCCCCTGGGCTCCGCCAAGAGCGCCAGCAGTGGACAAGCTGGTGTCC
AGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGTCCTGGGTTCCTGGACCGCGCCATGGCCGCCTCCAACTGGCCTCCGACTTCAA
TCCGGCATCCGCAAGGTGCTGTTCCTGGACGGCATCGACAAGGCCACGAGCACGAGCGCCTACCACCTGGCGGCCATGGCCGTCCCCGGCATCT
CCTGCCCCCCATCGTGGCCAAGGAGATCGTGGCCTCCTGCGACAAGTGCCAGCTGAAGGGCGAGGCCCTACGTGCTCCCCGGCGCCACCGGC
GGCAGCTGGACTGCACCCACCTGGAGGGCAAGATCATCCTGGACGCCCGTGGCCCGTGGCCCGTGAAGGTGATCCCCGACAACTTCACCCTGCCCGAGGC
CAGGAGACCGCCTACTTCATCCTGAAGCTGGCGGCCGCCAGGAGTTCGGCATCCCCTACAACCCCTACAACCAGTCCATGAACAAGGAGCTGAAGGC
CGCCTGCTGGTGGGCCAAGATCATCGCCGAGAGTTCGGCCTGAAGAACCCCAGTCCCGTGTTCATCCACAAGGGCAACAAGGAGCTGGCGGCTACTCC
TCGGCTGGGCGCGGAGCAGGCCGAGATGGCCAACATCCTGAAGAACCGCCGCCAGGCGCTGCAGAAGGAGCTGCAGAAGCTCGGCGGTGTACCGCGGA
GCCGGGCAGCGCGAGCATCCTGGAAGGCGCCCGCGACCTGTGTGGAGAGCGGCGCGACGACTGCGTGGTGCCCCGCC
CTCCCGCGCCAAGCCCATCTGGAAGGCCCGCTGGAGATGGCCGCGGCTACCGGCGGCGACGACTGCTGCCCGCCTCCCGCGCCAAGCATCAAGGTGGTGCCCCGCC
GCAAGGCCAAGATCATCCGCGACTACGGCAAGCAGATGGCCGGCGACGACTGCGTGGCCGGCGACGAGGACTAA

Fig. 126B

2003_CON_12_BF pol.OPT

```
TTCTTCCGGAGAACCTGGCCTTCCAGCAGGGCGAGGCCCGCAGCAGGCCCGGCGGCCAACTCCCCGCCTCCCGGGAGCTGTGGGTGCG
CCGGGGCGACAACCCCCTGTCCGAGGCCGGCGGCCGAGCGCCGGCCACCGTGCCTGCTGGAGGACCATCAACCTGCCCGGCAAGTGGAAGCCCAAG
CCATCAAGGTGGGGCGGCCAGCTGAAGCTGGCGGCTTCATCAAGGTGAAGCAGTACGACAACATCCTGATCGAGATCTGCGGCACAAGGCCATCGGCCACCGTGCTGCCGTGCTGGTGGGCCC
ATGATCGGCGGCCATCGGCGGCTTCATCAAGGTGAAGCAGTACGACAACATCCTGATCGAGATCTGCGGCACAAGGCCATCGGCACCGTGCTGCCGTGCTGGTGGGCCC
CACCCCGTGAACATCATCGGCCGCAACCTGCTGACCCAGCTGGGCTGCACCTGACCCTGAACTTCCCCATCTCCCCATCGAGACCTGAAGC
CCGGCATGGACGGCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGCAGGAAGATCAAGGCCCCTGACCGAGATCTGCACCGAGATGGAGAAGGAGGGCAAGATC
TCCAAGATCGGCCACGCACCCCTACAACACCCCCGTGTTCGCCATCAAGGAAGAAGGACTCCACCAAGTGGCGCAAGTTCCGCGAGCTGAA
CAAGGCCACCCAGGACTTCTGGGAGGTGCAGCTGGCGCATCCCCACCCCGCCTTCACCATCCCCGTGAACAACAGGAAGAGAAGAGAGTCCGTGACGTGGGCGACGCCT
ACTTCTCCGTGCCCCTGGAAGGGCTCCCCCCGCCATCTTCCAGTCCTCCTCCAGTACACCCCCAAGCAGAACCCGACATCGTGATCTACCA
CTGCCCCAGGGCTGGAAGGGCTCCCCCCGCCATCTTCCAGTCCTCCTCCAGTACACCCCCAAGCAGAACCCGACATCGTGATCTACCA
GTACATGGACGACCTGTACGTGGGCTCCGACCTGGAGATCGGCCAGCAGATCGAGGAGCTGCGCCAGCCACCTGACGGGCCTTCACCA
CCCCCGACAAGAGCACCAGAGAAGGAGCCCCCTTCCTGTGGATGGGCAAGCTGAACTGGGCCTGGGCTGCCGCCAAGGACTGGAACCTGAAG
GCGCGGCACCAAGGCCGTGAACGACCCTGAACGAGGTGATCCCGGCGGCCCAGCGAGTGTCCTGAAGGTGTCCTGAGTTCGTGA
TGTACTACGACCCCTCCAAGGACCTGATCGCCGGCGGGCCCACCAAACGACGTGAAGCTGGAAGCAGCCTTGGCCTCCCATCGTGATCGTGGGG
ACCGGCAAGTACGCCCCGTGGTGGACGAGATCGTGGCCCTGCCCATCGCCCTGCTCCAAGCTCTGAAGGAGCTGAACCGCCAGTTCGTGA
CAAGAGACCCCCCAAGTTCCGCCTGGTGCCCCATCCTGAAGGAGATGCCCGAGCTGCTGGAGCTGCCACTGACGTGGACGCGGCGCCATCCAGCTGGC
ACACCCCCCCTGGTGAACCGGCTACGTGGGCTCGGAGGTGTTGGACCGGCCATCCCGCCTGGGCATCGAGCTGCCGAGTCCGAGCTGGTGA
CCTGCAGGACTCCGGCCTACTTCATCCTGAGGTGAACATCGTGAAGGAGGTGAACATCGTGTACCTGTCCTGGGGTGCCCGCCAACAGCTGGTG
ACCAGATCATCGAGGCTGATCAAGAAGAAGAGGTGTACCTGCTGGACGGCATCGACAAGGGCCCAGGAGCACGAGAAGTACCACAACTGGCGCGCCATGGCCTCCGACTT
TCCGCCGGCATCCCCGTGTTCCTGGACGAGATCGTGGCCCAAGGAGATCGTGGCCTTCCTGGACCAAGAGCGAGGTGGAGTGGTGTGCCGTGTG
CAACCTGCCCCCCGTGGTGCCCAAGGAATCGTGGCCCCAAGATCATCTGGTTCGTGCTGAAGAGCGAGGTGCCATGGCCTGCTCCCCCGGCA
TCTGGCAGCTGGACTGCACCACCTTCACCTGGCCAAGGATCATCTGTCCGAGGGCCATGAGGTGATCCCGCCCGAGACC
GGCCAGGAGGAGCTGCCGCTACTTCATCCTGAGCTGAACGCCGCACCTGAACACCGACACCGAGGTCATCCTCCTCCGCGGCCGTGAA
GGCCGCCTGCTGGTGGGCGGAGCGGCCATCATCGACATCATCCAGAGCGGCCATCCAGGTGGAGTGTGCAGGCGGCGGCGCATCGGCCGGCTAC
TCATCCGCCCGGTGTGCGGCCGACCCCGAGGGGCCCCGACACTGGAAGGGCCGAGGGGGCCCGAGGTGGATCCCGCCGGCGCCGAGATCAAGGCGGGGCATCGGCCGGCTAC
CGACTCCGCCGACCCCGTGTCTGTGAAGGCCGACTACGGCGAAGGGCGGCCGACTACGGGGCGGCGGCGGACTAA
GCCGCAAGGCCAAGATCATCCGCGACTACGGCGAAGGGCGGCCGACTACGGGGCGGCGGCGGACTAA
```

Fig. 127A

84. 2003_CON_14_BG_pol.PEP
FFRENLAFQQGEAREFSPEQARANSPTRRELWVRRGDSPLPEARAEGKGDIPLSLPQITLWQRPLVTVRIGGQLIEALLDTGADDTVLEDIN
LPGKWKPKMIGGIGGFIKVRQYDQILIEICGKKAIGTVLVGPTPINIIGRNMLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEK
IKALTDICTEMEREGKISKIGPENPYNTPIFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPSGLKKKKSVTVLDVGDAYFSVPLDE
SFRKYTAFTIPSTNNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRIKNPEIVIYQYMDDLYVGSDLEIGQHRAKIEELRKHLLSWGF
TTPDKKHQKEPPFLWMGYELHPDKWTVQPIQLPDKESWTVNDIQKLVGKLNWASQIYPGIKVKQLCKLLRGAKALTDIVPLTAEAELELAEN
REILKEPVHGVYYEPSKELIAEVQKQGLDQWTYQIYQEPYKNLKTGKYAKRGSAHTNDVKQLTEVVQKIATESIVIWGKTPKFKLPIRKETW
EVWWTEYWQATWIPDWEFVNTPPLVKLWYRLETEPIAGAETYVVDGAANRETKLGKAGYVTDKGKQKIITLTETTNQKAELQAIHIALQDSG
SEVNIVTDSQYALGIIQAQPDRSESEVVNQIIEQLIKKEKVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGIDKAQEEHEKYHSNWRAM
ASDFNLPPVVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKIILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKIIH
TDNGSNFTSAAVKAACWWANITQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIAS
DIQTKELQKQITKIQNFRVYFRDSRDPIWKGPAKLLWKGEGAVVIQDNNEIKVVPRRKAKIIRDYGKQMAGDDCVAGRQDED$

Fig. 127B

2003_CON_14_BG pol.OPT

```
TTCTTCCGCGAGAACCTGGCCTTCCAGCAGGGCGAGGCCCGGCGAGTTCTCCCCCGAGCAGGCCCGGCGCCAACTCCCCCACCCGCCGCGCGAGCTGTGGGTGCG
CCGCGGCGACTCCCCCTGCCCGAGGCCCGGCGCCGAGGGCAAGGGCGACATCCCCTGTCCCGCCCAGATCACAACCTGTGGCAGCCCCCTGGTGACCG
TGCGCCATCGGCGGCCAGCTGATCGAGGGCTTCATCAAGGTGCGCGCAACAGATCCTGGAGACCTGAGATCGGCACGTGGTGCGTGGGCCCCAC
ATCGGGCGCATCGGCGGCTTCATCAAGGTGCGCGCAACATGTCGACCCAGATCGGCTGCACCCTGAACTTCCCCCATCTCCCCCTGAACCTGCCCGTGGGCGCTGAAGCTGGAAGCCCG
CCCATCAACATCATCGGCCGCAACATGTCGACCCAGATCGGCTGCACCCTGAACTTCCCCCATCTCCCCCTGAACCTGCCCGTGGGCCGTGAAGCTGGAAGCCCG
GCATGGACGGCCCCAAGGTGAAGCAGTAGTGGCCCCTGACCGAGAGAAGATCAAGGCCCTGACCGAGATGCTGCACCGAGATGGTGAGCTGAAGGGCAAGATCTCC
AAGATCGGCACCGAGAACCCCTACAACACCCCCATCTTCGCCATCAAGAAGAAGACTCCCGCCATCCCCCTCCCGGCTGCATCCCCCGAGCTGGCGGCCTACT
GCGCACCCCAGGACTTCTGGGAGGTGCAGCTGGGCTTCCCCGACGAGTCCTTCCCCGCCATCCCCCTCCCGGCTGCATCCCCCGAGCTGGCGGCCTACT
TCTCCGTGCCCCTGGACGAGTCCTTCCCCGCCATCTTCCATGACCAAGTACCACCCCTTCCCGCCATCAAGAACCCCGAGATCGTGATCTACCAGTA
CCCCAGGGCTGGAAGGGCTCCCCGACCTGGGCTCCAGAGATCGGCCAGCACCGGCCAAGATCGGCCAGCAGCACCTGTGCTGTCCCGACCACACCCC
CATGGACGACCTGTACGTGGGCTCCCGACCTGGGCTCCAGATGGGATGGGCTACGACGTGCACCCGACAGTGCAGCTGCAAGTGACCCCGACAAGGAG
CCGACAAGAAGCACCAGAAGCACCAGAACGACATCCAGAAGCTGGGCCTTCCTGTGGGCAAGCTGGGCTGAACTGGGCCGAGAGCTGTGAAGCTGCTGCG
TCCTGGACCGTGAACGACGACATCCAGAAGCTGGGCCTGTGCCCCTGACCGCCGAGCTGGCCCGAGAACCGCCGAGATCCTGAAGGTGAAGCAGCTGTGCAAGCGGCTGCG
CGGCGCCAAGGCCCTGACCGGCCGACATCGTGCCCGAGGTGCAGGGCGTGCAGGGTGCAGGAAGCAGGGCCCTACCAGGATCTACCAGGAGAACCGCCGAGAACCTGAAGACC
ACTACGAGCGCCCTCCAAGGAGCTGATCGCCGAGGTGATCGGCCGCTCCGCCCACACCGACGACGTGAAGCAGCTGATGGCAGCAGCTGGTGTGCAGCCCGGACCGCCAACCGGCAA
GGCAAGTACGGCCAAGCGCGGCTTCCGCCCACACCGACGACGTGGAAGCAGCTGATGGCAGCTGGTGTGGACCTGGACCTGGGGAGTTCGTGAACA
GACCCCCAAGTTCAAGCTGCCCATCCGCAAGGAGACCTGGGAGGTGTGTGGACCGGCGCCGAGGTGGCCTACTACGTGCCCGCGCCAACCGCGAGACCAAGCTG
CCCCCCCCGTGGTGAACGTGCCCAAGGCCTGGAACATCGTGACCGACTACGTGGGCTCCGGAGATCAGCGAGAGCGGCCCCGGCGCCGAGGTGATCCCCGCCGAGACCCGGC
GGCAAGGCCGGCTACGTGACCGACAAGGGCGACACCTGGAAGGTGTACCTGTCCGACGACAAGGCCCAGAAGTACCACTCCAGGCCCTCCGAGGTGGTGAACC
GCAGGACTCCGGTCCGAGGTGAACATCGTGACCGACTACGTGGGCTCCGGGAGAGCTGAGGCGCGGGCCAAGACGAGCAGGCCCAACGAGCAGGTGGTCC
AGATCATCGAGCAGCTGATCAAGAAGGAGAAAGGTGTACCTGTCCTGGACGGCATCGACAAGGCCCAGAAGTACCACTCCAGGCCGTCCGAGGTGGTGAACC
TCCCGGCATCGGCCAAGGTGCTGTTCCTGGACGGCATCGACAAGGCCCAGGAGGAGGCACGAGAAGTACCACCTCCAGCCCTCCGACTTCAA
CCTGCCCCCCGTGGTGACCGCTGGCCCTGGCCAAGATCGTGTGCCCTGACCGGCGCGCCAGCTGCACGTGGACTGATCCCCGCCATCT
GGCAGCTGACCTGCACCCACCTGGAGGGCAAGATCATCGTGGCCGTGCACGTGGCCTCCGGCTACATCGAGGCCGAGGTGATCCCCGCCGAGACCCGGC
CAGGAGAGACGCCGGCCTACTTCATCCTGAAGCTGGCCGGCCGTGTGGCCTGTCCTGGAGTCCCGCGGCCTCGACACGTGGCGTCC
CGCCTGCTGGTGGCGGCCATCAGGCCGAGCACCGCCGCCTACATCATCGGCCCTGAAGGGCGTGGGAGCTGATCCCCGCGCCGTGAAGGC
TCGGCCAGTGCGGCGGCCATCAGCGGCCAGATCCCCCGGCCCCCCCCAGTCCGTGCTGGCCAGATGGCCGGCTACTTCCGCGGCGTGTACTTCCGCGA
GCCCGGCCAGCTGGCGGCCATCAGCGGCCAGATCCCCCGGCCCCCCCCAGTCCCGTGCTGGCCAGATCAAGCGCGGCGGCTACTCCGGCGTGTACTTCCGCGA
CTCCCGCGACCCATCTGGAAGGCCCGGCTGCTGGTGCGGCCAAGCTGGCCAAGATCAAGGTGGTGCCCGGCC
GCAAGGCCAAGATCATCCGCGACTACGGCAAGCAGATGGCCGGCGACTGCGTGGCCGGCAGGGCGGAGGGCCGAGGCCAGGGAGGACTAA
```